The Dissertation Committee for Patricia Elena Salerno Certifies that this is the approved version of the following dissertation:

Evolutionary History of Lost World Frogs

Committee:

David C. Cannatella, Supervisor

David M. Hillis

Claus Wilke

Beryl B. Simpson

Francisco L. Perez

Evolutionary History of Lost World Frogs

by

Patricia Elena Salerno, BA

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

Dedication

I want to dedicate my dissertation to my incredible role models, who have guided me through life: my mother, my father, my grandmother, and my sisters. With a writer for a father, a university professor for a mother, and two older sisters who already hold their own PhDs, it's no wonder I've made it this far. My grandmother is there to remind us all of what matters the most: family. I love you all.

Acknowledgements

First and foremost, I need to thank my collaborators in Venezuela, Fernando Rojas-Runjaic and Celsa Señaris, without whom this project would have been impossible. Their support and guidance made a logistically infeasible project a reality, with much hard work, sweat, and tears. I would also like to thank my advisor David Cannatella, and my committee, David Hillis, Claus Wilke, Beryl Simpson, and Francisco Perez, for their support, guidance, and helpful advice throughout the way. I want to give enormous thanks to the "old generation" of the Cannatella-Hillis-Bull labs, who offered advice, support, and generous smiles during my first years in gradschool (Jeremy, Tracy, Rick, Shannon, Will, Thomas, Greg, Santiago, Beckie, Carlos, Jeanine), and Brian Caudle for being an amazing office mate and training me in the lab.

I would like to thank Fernando Rojas-Runjaic, Mark Dominguez, Greg Pauly, Cesar Barrio-Amorós, and the late Henry Viana for fieldwork assistance. I would like to thank A. Wynn, C. Barrio-Amorós, and B. Noonan for tissue sample loans. I also thank S. Hunicke-Smith and J. Podnar at the Genomic Sequencing and Analysis Facility (UT Austin), as well as D. Bolnick, J. Weber, and others in the Bolnick and Juenger labs (UT Austin) for generous help with the lab protocols, and access to reagents and equipment, without which the third chapter would not have been feasible. I would also like to thank T. Heath, M. Bernal, A. Wright, S. Maguire, B. Tarvin, E. Ortiz, and many others I'm surely forgetting for help with analyses and software.

Collecting and tissue usage permits were issued by the Venezuelan Ministerio del Poder Popular para el Ambiente, Comisión Nacional de Tepuyes, Instituto Nacional de Parques, and Comunidad Indígena de Kamarata. This work was approved by IACUC.

This work was funded by National Geographic/Waitt, NSF DDIG (DEB-1210035), University of Texas Faculty Research Grant, American Philosophical Society Lewis and Clark Fund for Exploration and Research, Explorer's Club Exploration Grant, and various grants from the Ecology, Evolution and Behavior program at the University of Texas at Austin.

Last but not least, I thank many who have supported me or accompanied me at different stages and in different ways (that were not already mentioned), particularly Maria José, Sofía, Laura, Alejandro B., Alejandro P., Nestor, Alexis, Mariana, Monica, Silvia, Walter, Oscar V., Oscar L., Vicky, Tak, Gustavo, Sandi, and also ask for forgiveness from anyone I forgot that rightly deserved to be on this list... seven years is a long time.

Thank you all.

Evolutionary History of Lost World Frogs

Patricia Elena Salerno, Ph.D.

The University of Texas at Austin, 2014

Supervisor: David C. Cannatella

The Lost World of South America is a unique landscape of flattop mountains that are home to hundreds of endemic species. These flattop mountains, or tepuis, were formed after millions of years of erosion of the high-altitude Guiana Shield plateau. The tepui summits, isolated by their surrounding cliffs that can be up to 1000 m tall, are thought of as "islands in the sky," harboring relict flora and fauna that underwent vicariant speciation due to plateau fragmentation. High endemicity atop tepui summits supports the idea of an ancient "Lost World" biota. However, recent work suggests dispersal between lowlands and summits occurred long after tepui formation, but neither view (i.e., ancient vicariance vs. recent dispersal) has strong empirical support owing to a lack of studies.

I tested diversification hypotheses of the Guiana Shield highlands by estimating divergence times of *Tepuihyla*, a Guiana Shield endemic genus. Diversification among the different species did not support the Lost World Hypothesis of summit diversification, but rather recent dispersal approximately 50 million years after tepuis formed. This study was the first to explicitly test these hypotheses with a tepui endemic

vertebrate, and as such a significant contribution to our understanding of the evolutionary history of this region.

After increasing sampling, I focused on three of the most recently diverged lineages of *Tepuihyla*, in order to examine population genetics, phylogeography, and species delimitation atop these summits. I found high levels of lineage sorting in spite of low divergences in both nuclear and mitochondrial genes. I also found an unexpected pattern of nuclear versus mitochondrial diversity, suggesting the possibility of a recent mitochondrial selective sweep. Species delimitation analyses support the existence of a cryptic, undescribed summit species.

Finally, I obtained a Single Nucleotide Polymorphism matrix with next-generation sequencing in order to observe more fine-scale population structure atop the Chimantá massif, a formation composed of ten tepui summits and intermediate altitudes separating them. I found high levels of population structure and assignment atop different tepui summits on the massif, indicating that even at extremely low levels of divergence the landscape complexity may be fomenting population isolation even at the smaller scale of within-massif divergences.

Table of Contents

List of Tables

List of Figures

Chapter 1: Ancient Tepui summits harbor young rather than old lineages of endemic frogs

ABSTRACT

The flattop mountains (tepuis) of South America are ancient remnants of the Precambrian Guiana Shield plateau. The tepui summits, isolated by their surrounding cliffs that can be up to 1000m tall, are thought of as 'islands in the sky,' harboring relict flora and fauna that underwent vicariant speciation due to plateau fragmentation. High endemicity atop tepui summits supports the idea of an ancient 'Lost World' biota. However, recent work suggests that dispersal between lowlands and summits has occurred long after tepui formation indicating that tepui summits may not be as isolated from the lowlands as researchers have long suggested. Neither view of the origin of the tepui biota (i.e., ancient vicariance vs. recent dispersal) has strong empirical support owing to a lack of studies. We test diversification hypotheses of the Guiana Shield highlands by estimating divergence times of an endemic group of treefrogs, *Tepuihyla*. We find that diversification of this group does not support an ancient origin for this taxon; instead, divergence times among the highland species are 2–5 Ma. Our data indicate that most highland speciation occurred during the Pliocene. Thus, this unparalleled landscape known as “The Lost World” is inhabited, in part, not by Early Tertiary relicts but neoendemics.[1]

[1] **Salerno PE**, Ron SR, Señaris JC, Rojas-Runjaic F, Noonan BP and Cannatella DC. 2012. Ancient tepui summits harbor young rather than old lineages of endemic frogs. *Evolution* 66(10): 3000–3013 – PES obtained samples and collected, performed analyses, and did most of the writing; SRR, JCS, FRR, and BPN contributed with samples, data collection, and writing; DCC contributed with analyses and writing.

INTRODUCTION

The tepuis (flattop mountains) of northern South America are arguably the most dramatic example of sky islands on Earth. With relatively flat summit plateaus sitting atop sheer cliffs of up to 1000m, these tabletop mountains form a discontinuous ecosystem called Pantepui (Huber, 1988) and harbor numerous endemic lineages. For example, within the Pantepui, 60% of the vascular plant species and 87% of the frog species are endemics, often to a single tepui summit (Berry and Riina, 2005; Duellman, 1999; McDiarmid and Donnelly, 2005). The tepuis are part of the Guiana Shield, which is situated between the Orinoco and Amazon Basins, accounts for 9% of the land area of South America, and shows biogeographic affinities with the Amazon and Andes.

Early reports of the unique topography and biodiversity of the Pantepui inspired Sir Arthur Conan Doyle to write of long-isolated dinosaur populations surviving to the present day in his novel "The Lost World," which inspired the namesake Lost World Hypothesis (Rull, 2004a; or Plateau Hypothesis, Chapman, 1931), which is the main biogeographic hypothesis put forward to explain the highly endemic fauna and flora of this region (Chapman, 1931; Maguire, 1970; Hoogmoed, 1979; McDiarmid and Donnelly, 2005; Heinecke et al., 2009). This hypothesis predicts that highland species, which are often endemic to a single tepui, are relicts of formerly widespread plateau taxa and arose through vicariance following ancient fragmentation of the plateau. Thus, summit taxa are hypothesized to have been isolated for millions of years. In examining the especially high endemism levels of Pantepui frog genera, most researchers support the idea that diversification must be explained, at least partially, by long-term isolation of Plateau

paleoendemics (Hoogmoed, 1979; MacCulloch and Lathrop, 2002; McDiarmid and Donnelly, 2005; Heinecke et al., 2009). However, to date it is unclear what fraction of the highland endemism, in frogs or other groups, is explained by the Lost World Hypothesis.

Many global biodiversity hotspots are tropical mountainous areas, suggesting that highland areas promote higher rates of diversification (Orme et al., 2005). In most uplifted mountain ranges such as the Andes, the summits are the most recently exposed surfaces, and these seem to be associated with recent species divergences (Hughes and Eastwood, 2006). In contrast, the tepui summits are of Precambrian origin and are thought to harbor ancient lineages (Chapman, 1931; Maguire, 1970; Hoogmoed, 1979; McDiarmid and Donnelly, 2005; Heinecke et al., 2009). The tepuis were formed about 70–90 Ma, after the Guiana Shield plateau underwent several periods of erosion and plateau uplift and fragmentation starting around 300 Ma, resulting in isolated sky islands (Briceño et al., 1990; Briceño and Schubert, 1990; Gibbs and Barron, 1993). Because the tepuis were formed mostly by erosion rather than uplift, the summit surfaces are geologically older than the adjacent slopes.

Many studies have suggested that vicariance alone (i.e., the Lost World Hypothesis) cannot explain the current distribution of highland vascular plants (Huber, 1988; Givnish et al., 1997), birds (Mayr and Phelps, 1969), and ants (Jaffe et al., 1993), nor the available pollen deposits (Rull, 2005; Rull and Nogué, 2007). A set of alternative hypotheses states that diversification occurred more recently, after the summits were colonized by lowland species through various mechanisms such as habitat shifts, vertical ("cool climate") displacement, and island hopping (Huber, 1988; Mayr and Phelps, 1967;

Rull, 2004a). The Island-Hopping Hypothesis (Chapman, 1931) suggests aerial dispersal among the tepuis. This hypothesis seems plausible (though not obligatory) for organisms with higher dispersal abilities (e.g. birds, vascular plants, insects; Mayr and Phelps, 1967; Jaffe et al., 1993; Givnish et al., 1997; Rull and Nogué, 2007) but unlikely for low vagility organisms. The Habitat Shift Hypothesis (Mayr and Phelps, 1967) suggests that lowland species adapted to cooler climates, allowing colonization of the highlands (Mayr and Phelps, 1967; Huber, 1988; McDiarmid and Donnelly, 2005). The Vertical Displacement Hypothesis (Rull, 2004a, 2004c, 2005), which is not exclusive of the Habitat Shift Hypothesis, suggests that cooler climates, especially during Quaternary climate oscillations, promoted downward elevational shifts of habitat and range expansion of highland species into the lowlands, thus connecting previously isolated populations. Subsequent warmer interglacials promoted upward shifts of these cold-adapted populations, isolating them on the summits. Thus, the Vertical Displacement Hypothesis specifically invokes historical climate change, whereas the Habitat Shift Hypothesis focuses on populations adapting to new habitats.

The Vertical Displacement Hypothesis is similar to the well-known Forest Refuge Hypothesis of Haffer (1969) in that both invoke glacial-interglacial climate fluctuations resulting in alternation between isolated refugia and widespread habitat types. Hypotheses of Pleistocene refugia, have been extremely popular, though also highly disputed, for explaining biogeographic patternsboth in temperate (Knowles, 2000; Johnson and Cicero, 2004; Galbreath et al., 2009) and tropical regions (Mayr and O'Hara, 1986; Hooghiemstra and Van der Hammen, 2004; Carnaval and Moritz, 2008). The

postulated Pleistocene refugia of the Guiana Shield are not restricted to tepui summits. For example, in low-elevation mountains in the eastern Guiana Shield, the effects of Pleistocene climatic fluctuation have been posited to promote extensive secondary contact among species typically found in mid-elevations (Noonan and Gaucher, 2005, 2006).

Debate on the origins of the Pantepui biota thus focuses on the relative importance of recent dispersal versus ancient vicariance. Not unexpectedly, the Pantepui biota is likely to be a mosaic of remnant Guiana Shield lineages and more recent colonizers (Rull, 2004b). Divergence times from molecular analyses of tepui species, and between these tepui inhabitants and their closest lowland relatives, can differentiate among these alternatives. To date, however, few studies have tested these hypotheses using phylogenetic approaches (but see Givnish et al., 1997, 2004), largely due to the great difficulty and expense of conducting fieldwork in the Pantepui ecosystem. To our knowledge, we offer the first explicit test of diversification hypotheses of a Pantepui-endemic vertebrate group by inferring the phylogenetic relationships and divergence times for species of *Tepuihyla*, which comprises seven allopatric treefrog species that occur only at mid- to high-elevations (Figure 1.1).

MATERIALS AND METHODS

Terminology

The flattop mountains that make up the Pantepui ecosystem are remnants of the Guiana Shield Plateau and derived from Roraima Group sandstones (Huber, 1987). The

Pantepui occupies mainly southern Venezuela but also adjacent regions of northeastern Guyana, southern Suriname, and northern Brazil (Huber, 1987). The concept of Pantepui that we follow is the association with the geological formation and not a particular elevation or floristic composition.

Samples and sequences

We sampled four of the seven recognized species of *Tepuihyla*. All *Tepuihyla* species are allopatric and inhabit montane and submontane areas. Of the seven species, four are known only from a single tepui, one (*T. edelcae*) from several adjacent tepui summits, and two (*T. rodriguezi* and *T. talbergae*) from lower elevation localities (Figure 1.1). One nominal species, *T. celsae*, is not considered in our discussions. It is known only from a single locality well outside the Pantepui region. Barrio-Amorós (2004) noted that the specimens of *T. celsae,* which were not collected by the authors who described the species, have a doubtful locality and he attributed this new species and locality to mislabeling. Although the poor condition of the specimens makes identification challenging, the specimen is likely a mislabeled *T. luteolabris* from Duida tepui (C. Barrio-Amorós, pers. comm.).

Tissue samples were obtained from the Pontificia Universidad Católica del Ecuador (QCAZ), Museo de Historia Natural La Salle, Venezuela (MHNLS), and United States National Museum (USNM). See Table 1.1 for GenBank accession numbers. The entire ingroup sample includes sequences of 86 individuals of Lophiohylini (Hylidae;

Trueb 1970; Faivovich et al. 2005), of which 43 were generated by us and 43 were downloaded from GenBank. The outgroup consists of 9 species of the sister clade Hylini.

Extraction and isolation of DNA, and amplification and sequencing of mitochondrial (mtDNA) and nuclear (nDNA) genes were done using standard techniques. Genomic DNA was extracted with the Viogene Blood and Tissue Genomic DNA Extraction Kit. Two different polymerase chain reaction (PCR) protocols were used, one for the nDNA (the pro-opiomelanocortin gene POMC, approx. 460bp) and another for the mtDNA (12S, valine tRNA, and 16S; approx. 2400bp). The *Tepuihyla aecii* tissue was from an old museum specimen, and extraction yielded low DNA concentrations, so only part of the 12S–16S could be amplified.

The mtDNA was amplified in four overlapping segments and using nine primers (MVZ59: 5'ATAGCACTGAAAAYGCTDAGATG3' [Goebel, 1999 #29], tRNA-phe: 5'GCRCTGAARATGCTGAGATGARCCC3' [Goebel, 1999 #30], 12LI: 5'AAAAAGCTTCAAACTGGGATTAGATACCCCACTAT3' [Goebel, 1999 #46], 12SM: 5'GGCAAGTCGTAACATGGTAAG3' [Pauly et al., 2004], tRNA-val: 5'GGTGTAAGCGAGAGGCTT3' [Goebel, 1999 #73], 16SH: 5'GCTAGACCATKATGCAAAAGGTA3' [Goebel, 1999 #76], 16SC: 5'GTRGGCCTAAAAGCAGCCAC3' [Pauly et al., 2004], 16SA: 5'ATGTTTTTGGTAAACAGGCG3' [Goebel, 1999 #87], 16SD: 5'CTCCGGTCTGAACTCAGATCACGTAG3' [Goebel, 1999]). The nDNA was amplified in one fragment (POMCR1: GGCRTTYTTGAAWAGAGTCATTAGWGG

[Vieites, 2007 #120], POMCF1: ATATGTCATGASCCAYTTYCGCTGGAA [Vieites 2007 #120]).

The thermocycler protocol for POMC was 2min at 94˚C followed by 12 cycles of 30sec@94˚C, 30sec@65˚C (–5˚C every 2 cycles), and 60sec@72˚C. The thermocycler protocol for the mtDNA was 2min at 94˚C followed by 34 cycles of 30sec@94˚C, 30sec@46˚C, and 60sec@72˚C. Amplified products were purified using the Viogene Gel-M Extraction Kit. Sequencing was performed using an ABI3100 PRISM sequencer.

Contiguous sequences were assembled for 12S-16S in Sequencher 4.8 and alignments were constructed using ClustalX 2.0 with default parameters (Thompson et al., 1997) followed by manual editing in MacClade 4.08 (Maddison and Maddison, 2002). Several regions in the mitochondrial dataset totaling 194 bases were judged to be unalignable and were excluded.

Phylogenetic analyses

The combined dataset of mtDNA and POMC included 3039 bases. All individuals were represented in the mtDNA alignment, but sequences of POMC were not available for all individuals so three analyses were performed. The first included 84 ingroup individuals plus outgroups (95 individuals total) with complete mtDNA data. To minimize missing data, the second included 46 individuals plus 7 outgroup species with complete data for mtDNA and POMC (53 individuals). The third included only POMC sequences (53 individuals), so that the nDNA and mtDNA topologies could be compared.

Three partitioning schemes were used in MrBayes (Ronquist and Huelsenbeck, 2003): a single partition including all data; two partitions, by locus (mtDNA and POMC); and three partitions, by gene and codon position (separating third codon position for POMC). Bayes factors were used to determine the preferred partition (Kass and Raftery, 1995), which was by gene and codon position. This partition was used in all subsequent analyses. The best-fitting model of evolution for each of the four partitions, GTR + G + I, was determined using MRMODELTEST (Nylander, 2004). Maximum likelihood analyses were performed using RAXML using the model GTR + G rather than GTR + G + I (Stamatakis, 2006). Node support was measured by standard nonparametric bootstrapping (1000 replicates) using RAXML, and by Bayesian posterior probabilities using MRBAYES (2 independent runs, 10 million generations, sampled every 1000, with burn-in of 1000 out of 10,000 samples). Convergence and stationarity of runs was assessed in TRACER (Drummond and Rambaut, 2007).

We compared the results for Tepuihyla to another frog group, *Stefania*, with similar Pantepui distribution. *Stefania* is a Pantepui endemic clade, but only two of the species available on GenBank occur on tepui summits. We added all available Genbank data for 12S and 16S (1510bp total) for 5 of the 19 described species of *Stefania* (Hylidae). The sequences included were the following: *Flectonotus fitzgeraldi* (AY819355, DQ679381), *Stefania coxi* (DQ679265, DQ679415), *Stefania ginesi* (DQ679266, DQ679417), *Stefania scalae* (DQ679267, DQ679418), *Stefania schuberti* (AY843768), *Stefania evansi* (AY843767), and *Stefania evansi* (AY819359). The

sequences were aligned and analyzed using the same protocols and programs as for *Tepuihyla*.

Divergence time estimates

Divergence times were estimated using BEAST (Drummond and Rambaut, 2007), which allows for simultaneous inference of topology and divergence times. We used two general methods of calibration to test for consistency among divergence estimates: fossils and paleobiogeography, and published substitution rates. The fossil calibrations included five hylid frogs (Smith et al., 2005). These are problematic for several reasons. First, only one pre-Holocene fossil is assigned to Lophiohylini (*Osteopilus septentrionalis* from the Pleistocene of the Bahamas; Sanchíz, 1998) and none for the group under study (*Osteocephalus* and *Tepuihyla*). Second, for many frog fossils the only recovered elements are ilia; rarely are even partially articulated frogs found, rendering identification less certain. Third, assignment of frog fossils to species or genera is often done from overall similarity rather than evidence of phylogenetic relationship from synapomorphies (Bell et al., 2010). Nonetheless, dates of occurrence of some frog fossils, assigned only on the basis of general similarity (Holman, 1998; 2003) have been used to calibrate chronograms of hylid frogs (Smith et al., 2005; Lemmon et al., 2007).

We also used a paleobiogeographic calibration to the common ancestor of *Osteopilus* and *Phyllodytes auratus*, using the GAARlandia or Aves Ridge landbridge hypothesis (Iturralde-Vinent and MacPhee, 1999). GAARlandia is a hypothesized Caribbean landbridge that connected the Greater Antilles to northern South America

between 33 and 35 Ma. The support for this hypothesis is mixed; some authors support it as a major means of dispersal to the Caribbean islands (i.e. mammals, spiders, frogs; Davalos, 2004; Crews and Gillespie, 2010; Alonso et al., 2012) while others strongly oppose it (Hedges, 2006). However, Moen and Wiens (2009) found that the divergence of *Osteopilus* from other mainland Lophiohylini overlapped with the 33–35 Ma timeframe of this hypothesis. The prior for the GAARlandia node was drawn from a normal distribution with a mean of 34 Ma and standard deviation of 1 Ma, based on the proposed range of 33–35 Ma (Iturralde-Vinent and MacPhee, 1999). Although the GAARlandia hypothesis is highly debated, we believe the inclusion of this calibration did not bias our results, since it did not change overall estimates. Furthermore, it substantially improved the computational performance of the analysis, in that good effective sample sizes (ESS > 200) were never reached (>80M generations) without the GAARlandia calibration. With the calibrations we analyzed the mtDNA dataset only as well as the combined mtDNA and nDNA datasets. Data partitioning was identical to that in the MRBAYES analysis.

We compared chronograms obtained from fossil and paleobiogeographic calibrations to chronograms generated using two estimates of rates of evolution for anuran mitochondrial 12S and 16S genes, one from *Xenopus* (0.00249 substitutions per site per lineage per Myr, hereafter the "*Xenopus* rate"; Evans et al., 2004) and one from *Pseudacris* (0.00277 substitutions per site per lineage per Myr, the *"Pseudacris* rate"; Lemmon et al., 2007). The *Pseudacris* rate also relied in part on the hylid frog fossil calibrations (Smith et al., 2005) so it is not independent of our rate estimates. For these

analyses only the mitochondrial dataset was used because the published rates of evolution (Evans et al., 2004; Lemmon et al., 2007) were not estimated using POMC.

The BEAST analyses were run with sufficient generations (35–60 million) to yield effective sample sizes of at least 200 for all parameters. The GTR + G model was used with uncorrelated lognormal distributions of branch lengths and with no specification of a prior tree. Every 1000th generation was sampled, and the first 10% of the samples were discarded as burn-in based on examination of all parameter estimates in TRACER. Trees were summarized using TREEANNOTATOR (in the BEAST package) with the target tree type set as maximum clade credibility.

The analyses of the *Stefania* Genbank sequences were conducted similarly as in *Tepuihyla/Osteocephalus.* Two separate analyses were performed using the *Xenopus* and *Pseudacris* rates. Given that the *Stefania* dataset is small (compared to *Tepuihyla*), we only used BEAST (and exclude RAxML and MrBayes analyses), to estimate tree topology and divergence times.

RESULTS

Phylogenetic analyses

Maximum likelihood and Bayesian analyses indicate that species of *Tepuihyla* (4 of 7 recognized species; 11 individuals) form a strongly supported clade for both mitochondrial and nuclear data (1.0 Bayesian posterior probability, BPP; Figures 1.2 and 1.3; 100% likelihood bootstrap support, BS; Figure 3.3). The samples of *Osteocephalus*

(11 of 24 species) revealed that *Tepuihyla* and *Osteocephalus exophthalmus* form a strongly supported clade (100% BS, 1.0 BPP; Figures 1.2 and 1.3) that is the sister group of all other *Osteocephalus*. Thus *Osteocephalus* as currently delimited is paraphyletic, and henceforth, we denote this by referring to "*Osteocephalus" exophthalmus* with quotation marks pending our resolution of this taxonomic issue, which is in progress.

We also analyzed all genera within the Lophiohylini, which includes *Osteocephalus* and *Tepuihyla*. Although the Lophiohylini was strongly supported as monophyletic, relationships among the deeper nodes of the lophiohyline clades (Figure 1.3) are generally poorly supported and our topology is somewhat inconsistent with other studies (Salducci et al., 2002; Faivovich et al., 2005; Wiens et al., 2010). However, the sister-group relationship between *Tepuihyla* and *Osteocephalus* holds in all studies.

The divergences among the *T. edelcae* samples from the summits of Auyan Tepui and the Chimantá Massif (Figure 1.1) are similar to the among-species divergences within *Tepuihyla.* Further, the mtDNA and nDNA phylogenies are incongruent regarding the monophyly of *T. edelcae*. The mtDNA only (Figure 1.2) and the combined mtDNA and POMC (Figure 1.3a) topologies both group *T. edelcae* from Chimantá as more closely related to the mid-elevation *T. rodriguezi/T. talbergae* than to the *T. edelcae* from Auyan indicating that *T. edelcae* is paraphyletic. However, based on nDNA only (Figure 1.3b), *T. edelcae* is monophyletic, though this is poorly supported (44 BS, 0.81 BPP). The low variability in the nuclear locus does not resolve the species-level relationships within *Tepuihyla*. Thus, the apparent discrepancy between the nuclear and mitochondrial

trees for *T. edelcae* might be due to mitochondrial introgression from recent hybridization between *T. edelcae* and *T. rodriguezi/T. talbergae* or to incomplete lineage sorting.

Divergence time estimates

Divergence time estimates based on calibrations from either fossils or substitution rates showed similar results (Figure 1.4). Taking into account all results from different datasets and calibrations, divergences between the Pantepui-endemic *Tepuihyla* and its sister taxon were estimated to be 14.7–23.6 Ma, which is far more recent than the formation of the tepuis. Furthermore, the oldest estimates for the divergence between this clade (*Tepuihyla* + *"O." exophthalmus*) and *Osteocephalus* are still at least 20 Ma more recent than tepui formation. Divergences among *Tepuihyla* species or populations inhabiting different tepuis are also relatively recent, from 0.7 to 8.1 Ma, with several divergences overlapping the Quaternary.

The divergence times from the *Xenopus* estimate were the oldest, possibly because *Xenopus* is very distantly related. The rates found using the hylid fossil calibrations are all in close agreement, the only difference being that the estimates obtained from the fossil calibrations (compared to the *Pseudacris* estimate) had much higher variance, probably because none of the calibrations lies within the group of interest (*Tepuihyla* and *Osteocephalus*). The only calibration within Lophiohylini is the paleobiogeographic calibration of *Osteopilus*. Given this, we focus the discussion on the analysis using the *Pseudacris* estimate. Estimated divergence times among *Tepuihyla*

species range from 0.7–5.3 Ma. The oldest node, the split between *T. aecii* and all other *Tepuihyla*, is estimated at 5.3 Ma (Figure 1.2, node A). Notably, this split corresponds to the geographic division between the western tepui group and the eastern tepui group (Huber, 1988). *Tepuihyla aecii* is from Cerro Duida in the western group, and the remaining species are from the eastern group.

The youngest split (0.7 Ma) is between *T. talbergae* (Kaieteur National Park, Guyana, 366m, but known from higher elevations) and *T. rodriguezi* (Gran Sabana, Venezuela, 800–1200m); both species inhabit the lowest elevations known for *Tepuihyla*. Not all analyses recovered this poorly supported node; in the mtDNA tree (Figure 1.2) *T. rodriguezi* was paraphyletic, which suggests either incomplete lineage sorting, or simply that they are a single species. We refer to this complex hereafter as *T. talbergae/rodriguezi.* The divergence between *Tepuihyla* and its sister species *"O." exophthalmus* is estimated at 14.7 Ma (Figure 1.2; node B), and that between *Tepuihyla* + *"O." exophthalmus* and *Osteocephalus* is estimated at 24.7 Ma (Figure 1.2; node C).

The results of the *Stefania* analyses using the *Xenopus* and *Pseudacris* rates were very similar, so we only report the divergence times from the *Pseudacris* rate. The deepest divergence found is 34.4 Ma, between *S. ginesi* and all other included species (Figure 1.5). The most recent divergence is estimated at 7.1 Ma, between *S. scalae* and *S. evansi*.

DISCUSSION

Monophyly and divergences among* Tepuihyla *species

The low divergence estimates among *Tepuihyla* species (<5.3 Ma) indicate that this Pantepui clade did not speciate under a Lost World vicariance scenario. Even when accounting for uncertainty in rates of molecular evolution, credibility intervals, and time calibrations, it is evident that species diversity within this clade is not the product of ancient dissection of the Guiana Shield plateau. Furthermore, the origin of the Pantepui-endemic treefrogs does not predate the formation of the tepuis, which is another prediction of the Lost World Hypothesis.

Previous phylogenetic analyses that included lophiohyline taxa (Faivovich et al., 2005; Wiens et al., 2010) sampled only one species of *Tepuihyla* so the monophyly of this group had never been tested. With our increased sampling (4 of 7 recognized species), we found *Tepuihyla* to be monophyletic. However, *Osteocephalus* as currently delimited (Faivovich et al., 2005; Trueb, 1970) is not monophyletic because "*O.*" *exophthalmus* was found to be the sister taxon to *Tepuihyla*. "*Osteocephalus" exophthalmus* is a Pantepui-endemic known from low and middle elevations in Guyana. In contrast to this clade of Guiana shield endemics (i.e., *"O." exophthalmus* + *Tepuihyla*), *Osteocephalus* is primarily a lowland Amazonian clade; a few species occur in the lowlands of the Guiana Shield. Thus, the key node to understanding the divergence of the Pantepui clade is that between *Tepuihyla* + '"*O.*" *exophthalmus* and *Osteocephalus* (node 6, Figure 1.4).

Tepuihyla + "*O.*" *exophthalmus* diverged from *Osteocephalus* about 24.7 Ma (Figure 1.2 and Figure 1.4, node 6). This timing of separation between *Osteocephalus* and *Tepuihyla* overlaps with that of a marine incursion into the present Amazon Basin

starting in the Early Miocene, 23–16 Ma (Hoorn, 1993). This may explain the largely allopatric distributions of *Osteocephalus* and *Tepuihyla*.

Other biogeographic hypotheses

Having rejected the Lost World Hypothesis for *Tepuihyla*, it is now clear that the divergence times among *Tepuihyla* species require dispersal between the lowlands and the summits. All three dispersal hypotheses (Island-Hopping, Habitat Shift, and Vertical Displacement) generate predictions related to dispersal frequencies and times of divergence. The Island-Hopping hypothesis is unlikely because island hopping should be a rare event for *Tepuihyla* given its low vagility. In addition, our finding of multiple relatively recent dispersal events is inconsistent with the predictions of Island Hopping. However, given our estimates of divergence times and phylogenetic relationships, the biogeographic history of *Tepuihyla* does not definitively reject either of the remaining dispersal hypotheses.

Given the occurrence of *Tepuihyla* in mostly middle and high elevations, as well as the recent divergence times, the establishment of the current distribution probably involved elements of both the Habitat Shift and Vertical Displacement hypotheses. Because all known *Tepuihyla* species are associated with the Pantepui, we infer that the common ancestor of *Tepuihyla* was also a Pantepui inhabitant that adapted to these conditions (as the Habitat Shift Hypothesis suggests). However, the sympatry of "*O." exophthalmus* with *T. rodriguezi /T. talbergae* throughout their limited distributions in

lower Pantepui elevations of the eastern Guiana Shield suggests that the association of this clade ("*O.*" *exophthalmus* + *Tepuihyla*) with the tepuis and possible adaptation to higher elevations occurred as far back as 14.7 Ma (Figure 1.2).

It is unlikely that Pleistocene climatic shifts were the only processes involved in divergences within *Tepuihyla*. The three deepest divergences within *Tepuihyla* pre-date the Quaternary, supporting the overlooked importance of pre-Quaternary diversification (Rull 2008). On the other hand, large confidence intervals among *Tepuihyla* divergences that overlap part of the Pleistocene, plus the small divergences between *T. edelcae* and *T. rodriguezi/T. talbergae* suggest that Pleistocene glacial cycles may have shaped the current distributions of these species. However, to determine the general significance of Pleistocene climate fluctuations, extensive sampling within recently diverged *Tepuihyla* species is needed to estimate recent gene flow using coalescent analyses.

Comparisons with phylogenies of other Tepui taxa

This analysis is the first to estimate DNA-based divergence times of a Pantepui-endemic animal taxon. Our analyses of *Tepuihyla* recover relatively recently diverged species (Plio-Pleistocene) with good species sampling (4/7 species). Although small samples of two other Pantepui frog groups have been analyzed phylogenetically (*Stefania* and *Ceuthomantis*, see below), the only other studies (to our knowledge) that explicitly treat the historical biogeography of a Pantepui clade include members of the Bromeliaceae and Rapateaceae (Givnish et al., 1997, 2000, 2004, 2011).

Bromeliaceae and Rapateaceae are families of flowering plants within the order Poales. Within the Rapataceae, the tepui-endemic crown clade Stegolepidieae (*Stegolepis*, *Amphiphyllum*, and *Epidryos*) diversified 10 Ma (Givnish et al., 2004). Dispersal, and not vicariance, is argued to be the principal correlate of divergence among these lineages (Givnish et al., 2004). Furthermore, the presence of some *Stegolepis* in lowland as well as intermediate and summit habitats corroborates recent dispersal.

Similarly, divergence times for many clades within Bromeliaceae have been estimated. The most recent common ancestor of crown-group Bromeliaceae likely inhabited the Pantepui at 19.1 Ma (Givnish et al., 2011, Figure 1.7). Brocchinioideae and Lindmanioideae, the two earliest-branching lineages within Bromeliaceae, are tepui endemics. The range of divergences between these clades is 8.9–19.0 Ma. The Bromeliaceae, then, does not owe its diversification (as a crown group) to the vicariant dissection of Pantepui, but more likely to late Cenozoic dispersal across a dissected landscape, a scenario similar to that observed for *Tepuihyla*.

The Pantepui endemic bromeliads in the genus *Brocchinia* are incredibly diverse in morphology and ecology (Givnish et al., 1997). This diversity could suggest that this clade is an "ancient" lineage. However, species within *Brocchinia* diverged relatively recently; the age of the crown group is at least 13.1 Ma (Givnish et al., 2001; Figure 1.7). Interestingly, *Tepuihyla* species seem to depend on water accumulation in phytotelmata of some species of *Brocchinia* (*B. hectiodes* and *B. acuminata*; Ayarzagüena et al., 1992), but no hypothesis of co-divergence of *Tepuihyla* and *Brocchinia* has been put forward.

We also examined divergence times of another Pantepui-endemic group, *Stefania*, for comparison to the *Tepuihyla* results. *Stefania*, like *Tepuihyla,* are highly dependent on the phytotelmata of pitcher plants. However, the reproductive biology of *Stefania* is quite distinct in that females are obligate dorsal egg-brooders and the embryos undergo direct development (Salerno and Pauly, 2012). In contrast, *Tepuihyla* have the more common reproductive mode of depositing eggs in bodies of water with the eggs hatching into a larval (tadpole) stage that undergoes metamorphosis (Ayarzaguena et al., 1992). Several authors have suggested that distributions and phylogenetic relationships in *Stefania* support a vicariant speciation scenario following the Lost World hypothesis (Hoogmoed, 1979; MacCulloch and Lathrop, 2002; McDiarmid and Donnelly, 2005). However, as is the case for most Pantepui groups, these assessments were based on limited sampling and distributional data with relationships among individuals solely based on morphological comparisons. Our analysis of the available sequences (5 of 19 species) shows that the deepest divergence within *Stefania* is 34.4 Ma, and the youngest is 7.1 (Figure 1.5). Interestingly, the youngest divergence is found between two lowland/midland species (*S. scalae* and *S. evansi*), which is the same pattern, though different timing, observed for the two *Tepuihyla* lowland/midland clades, *T. rodriguezi/T. talbergae*. Furthermore, these two species pairs have similar geographic distributions, which may indicate shared biogeographic histories. Even though the divergence estimates for *Stefania* clades are on average more deeply diverged than *Tepuihyla*, Stegolepidae, *Brocchinia*, and *Lindmania*, the divergences within *Stefania* indicate that this lineage radiated more recently than the

dissection of the tepui summits (70–90 Ma). Thus, *Stefania* follows the general diversification pattern in *Tepuihyla* and Bromeliaceae.

Many endemic representatives of the 'Lost World' are touted as "living fossils," that is, a species or group of species that was formerly speciose or widespread in time and/or space, but has suffered extinction (Brown and Lomolino, 1998). However, highly endemic clades are not necessarily "living fossils." There is no evidence that the clades *Tepuihyla*, *Stefania*, *Lindmania*, *Brocchinia*, and Stegolepidae are remnants of ancient widespread lineages. These taxa are more appropriately termed neoendemics, i.e., endemic taxa of relatively recent autochthonous origin (Brown and Lomolino, 1998).

The frog *Ceuthomantis smaragdinus* from Mt. Kopinang, Guyana also is claimed to be a "living fossil" (Heinecke et al., 2009). In contrast to the previous examples, this species is deeply diverged from its extremely speciose sister taxon at about 60 Ma (Heinecke et al., 2009). Two other species are tentatively referred to this genus, but no sequences are available. Although this divergence time is much greater than that of the other taxa considered here, the divergence of only one species makes it difficult to determine the minimal age at which that lineage (one species) was present within the Pantepui. Thus, further work is needed to assess whether this lineage is a paleoendemic.

Tepuis as current physical and ecological barriers

Many tepuis have sheer cliffs that have been proposed to be physical barriers to dispersal between low and high elevations, thus isolating summit taxa (Chapman, 1931;

Maguire, 1970; Hoogmoed, 1979). However, divergences among *Tepuihyla* species took place relatively recently, at least 60 Ma after the tepuis were fully formed, indicating that tepui walls are not a complete physical barrier between the summits and lowland forest even for organisms with low vagility.

Tepuihyla has an extremely fragmented distribution, but it is widespread in the Guiana Shield, occurring in three of the four Tepui provinces (Huber, 1988). All species are endemic either to the summits or to mid-elevation regions of tepui remnants and are generally found in Rapateaceae or Bromeliaceae-dominated meadows (Ayarzaguena et al., 1992). That *Tepuihyla* is absent from the lowlands between the tepuis suggests the habitat is currently unsuitable. Physiological constraints associated with increasing elevation (and thus decreasing temperature and increasing daily and yearly temperature variation) have been shown to enhance isolation among populations at different elevations, especially in ectothermic vertebrates (Navas, 2006). The drastic elevational differences between tepui summits and the surrounding lowlands result in extreme differences in annual mean temperature, precipitation, soil composition, and phytogeographic regions (Steyermark, 1979; Huber, 2006). Drastic microclimatic and habitat differences between summits and lowlands may create a challenge for dispersal, thus restricting *Tepuihyla* to suitable mid- and high-elevation conditions. However, as stated in the Vertical Displacement hypothesis, these climatic conditions may have been different in the past, allowing for a more widespread lowland existence.

Our results indicate that the sheer escarpments of tepuis have not prevented dispersal of *Tepuihyla* species across the lowlands between the summits during the last

5.3 Ma. Furthermore, the first association of the ancestor of *Tepuihyla* + "*O*." *exophthalmus* with the Pantepui likely occurred around 14.7 Ma, which deeply post-dates tepui fragmentation. Because most *Tepuihyla* divergences are during the Pliocene (5.3–2.6 Ma), our results also highlight the importance of pre-Quaternary speciation (Rull, 2008). However, we cannot completely reject the effects of Pleistocene glaciation, because the timing of the most recent divergences is consistent with climatic fluctuations. Our analysis clearly demonstrates dispersal of *Tepuihyla* to and/or from the tepui summits long after their formation. Furthermore, comparisons to other taxa such as *Stefania* and Bromeliaceae seem to indicate that dispersal has occurred across widely different tepui taxa, from pollen-dispersing plants to low-vagility organisms such as frogs. Thus, even though the Lost World hypothesis is attractive in nature and has been largely popular in the literature, so far there is no empirical dataset that shows unambiguous support for it.

ACKNOWLEDGEMENTS

We thank A. Wynn (United States National Museum) for tissue sample loans, the Ministerio del Poder Popular para el Ambiente (#2373, #0076, #41-0088) and Inparques (#0271) for permits, and C. Barrio-Amorós, G. Pauly, and members of the Cannatella lab for comments on the manuscript. This work was supported by funds from the American Philosophical Society, Explorers Club Exploration Grant, and the Ecology, Evolution and Behavior Graduate Program of the University of Texas to P.E.S., and by U. S. National Foundation Grant EF0334952 to D.C.C. and D.M. Hillis.

Table 1.1: List of specimens. Specimens that were sequenced are bold and marked with an asterisk. All other specimens were obtained from GenBank. Field/Museum codes are only shown for specimens sequenced herein.

	Taxon	Field/Museum code	Genbank Accession #s	
			12S-16S	POMC
1	*Acris crepitans*	-	EF566969	AY819109
2	*Aparasphenodon brunoi*	-	AY843567	-
3	*Argenteohyla siemersi*	-	AY843570	-
4	*Corythomantis greeningi*	-	AY843578	-
5	*Hyla arborea*	-	AY843601	DQ57787
6	*Hyla arenicolor*	-	EF566960	-
7	*Hyla cinerea*	-	AY680271	AY819116
8	*Hyla gratiosa*	-	AY843630	GQ374915
9	*Hyla meridionalis*	-	EF566953	GQ374915
10	*Hyla squirella*	-	AY843678	AY819120
11	*Hyla versicolor*	-	EF566953	DQ55805
12	*Itapotihyla langsdorffii*	-	AY843706	AY843706
13	***Itapotihyla langsdorffii****	USNM303287	JQ686500	JQ868470
14	*Nyctimantis rugiceps*	-	AY843781	-
15	*Osteocephalus alboguttatus*	-	DQ380347	-
16	*Osteocephalus buckleyi*	-	DQ380378	EUO34116
17	*Osteocephalus cabrerai*	-	AY843705	-
18	***Osteocephalus deridens****	QCAZ20868	JQ868501	JQ868484
19	***Osteocephalus exophthalmus****	BPN166	JQ868523	-
20	***Osteocephalus exophthalmus****	MHNLS19584	JQ868525	JQ868483
21	***Osteocephalus exophthalmus****	MHNLS19583	JQ868524	-
22	***Osteocephalus fuscifacies****	QCAZ20790	JQ868502	-
23	***Osteocephalus fuscifacies****	QCAZ20788	JQ868503	JQ868499
24	*Osteocephalus leprieurii*	-	AY843707	-
25	*Osteocephalus leprieurii*	-	AY549361	-
26	***Osteocephalus leprieurii****	MHNLS18689	JQ868505	JQ868497
27	***Osteocephalus leprieurii****	MHNLS18619	JQ868504	JQ868498
28	*Osteocephalus mutabor*	-	DQ380379	EUO34117
29	*Osteocephalus oophagus*	-	AF467267	-
30	*Osteocephalus oophagus*	-	AY843708	-
31	*Osteocephalus planiceps*	-	DQ380380	EUO43118
32	***Osteocephalus planiceps****	QCAZ19195	JQ868521	JQ868495
33	***Osteocephalus planiceps****	QCAZ20797	JQ868522	JQ868494
34	***Osteocephalus planiceps****	QCAZ18844	JQ868520	JQ868496

Table 1.1 *continued*

35	***Osteocephalus alboguttatus****	QCAZ18186	JQ868516	JQ868493
36	***Osteocephalus sp****	QCAZ38420	JQ868526	-
37	***Osteocephalus cf. taurinus****	USNM302469	JQ868514	-
38	***Osteocephalus cf. taurinus****	USNMFS008803	JQ868515	-
39	***Osteocephalus cf. taurinus****	MHNLS18325	JQ868506	-
40	*Osteocephalus taurinus*	-	AY843709	-
41	***Osteocephalus taurinus****	MHNLS18663	JQ868509	JQ868490
42	***Osteocephalus taurinus****	MHNLS15622	JQ868507	JQ868492
43	***Osteocephalus taurinus****	PS004	JQ868512	JQ868487
44	***Osteocephalus taurinus****	MHNLS19633	JQ868511	-
45	***Osteocephalus taurinus****	QCAZ18839	JQ868513	JQ868488
46	***Osteocephalus taurinus****	MHNLS17336	JQ868508	JQ868491
47	***Osteocephalus taurinus****	MHNLS18715	JQ868510	JQ868489
48	*Osteocephalus taurinus*	-	AY326041	AY819130
49	***Osteocephalus verruciger****	QCAZ13225	JQ868517	JQ868486
50	***Osteocephalus verruciger****	QCAZ17283	JQ868518	-
51	*Osteocephalus verruciger*	-	DQ380381	-
52	***Osteocephalus yasuni****	QCAZ19245	JQ868519	JQ868485
53	*Osteopilus crucialis*	-	AY843710	EUO34121
54	*Osteopilus dominicensis*	-	AY843711	-
55	*Osteopilus dominicensis*	-	AY819443	EUO34122
56	*Osteopilus marianae*	-	DQ380383	EUO34123
57	*Osteopilus pulchrilineatus*	-	AY819436	EUO34124
58	*Osteopilus septentrionalis*	-	AY843712	AY819131
59	*Osteopilus vastus*	-	AY843713	EUO34128
60	*Osteopilus wilderi*	-	DQ380385	EUO34129
61	*Phyllodytes auratus*	-	AY819383	AY819133
62	*Phyllodytes luteolus*	-	AY843721	-
63	*Phyllodytes sp*	-	AY843722	-
64	*Pseudacris crucifer*	-	AY291103	EF988269
65	***Tepuihyla aecii****	MHNLS12013	JQ868533	JQ868478
66	*Tepuihyla edelcae*	-	AY843770	-
67	***Tepuihyla edelcae****	PS002	JQ868537	JQ868475
68	***Tepuihyla edelcae****	MHNLS16090	JQ868534	JQ868477
69	***Tepuihyla edelcae****	MHNLS05824	JQ868535	-
70	***Tepuihyla edelcae****	PS001	JQ868536	JQ868476
71	***Tepuihyla edelcae****	PS268	JQ868538	-
72	***Tepuihyla rodriguezi****	PS003	JQ868540	JQ868474
73	***Tepuihyla rodriguezi****	MHNLS19575	JQ868539	-
74	*Tepuihyla sp*	-	DQ380389	EUO34131
75	***Tepuihyla talbergae****	BPN1101	JQ868541	-

Table 1.1 *continued*

76	***Tepuihyla talbergae****	BPN1219	JQ868542	JQ868473
77	*Trachycephalus coriaceus*	-	DQ380386	EUO34130
78	*Trachycephalus hadroceps*	-	AY843717	-
79	***Trachycephalus jordani****	QCAZ17509	JQ868527	JQ868471
80	*Trachycephalus jordani*	-	AY819395	AY819145
81	*Trachycephalus jordani*	-	AY326042	-
82	*Trachycephalus jordani*	-	AY843771	-
83	*Trachycephalus mesophaeus*	-	AY843718	-
84	*Trachycephalus nigromaculatus*	-	AY843772	-
85	***Trachycephalus resinifictrix****	QCAZ20808	JQ868528	JQ868481
86	*Trachycephalus resinifictrix*	-	AY843719	-
87	***Trachycephalus resinifictrix****	QCAZ19304	JQ868529	JQ868482
88	***Trachycephalus sp****	QCAZ21282	JQ868530	JQ868479
89	***Trachycephalus venulosus****	QCAZ21283	JQ868531	JQ868480
90	***Trachycephalus venulosus****	PS013	JQ868532	JQ868472
91	*Trachycephalus venulosus*	-	AY549362	-
92	*Trachycephalus venulosus*	-	AY326048	-
93	*Trachycephalus venulosus*	-	AY819382	-
94	*Trachycephalus venulosus*	-	DQ347027	-
95	*Trachycephalus venulosus*	-	AY364350	-

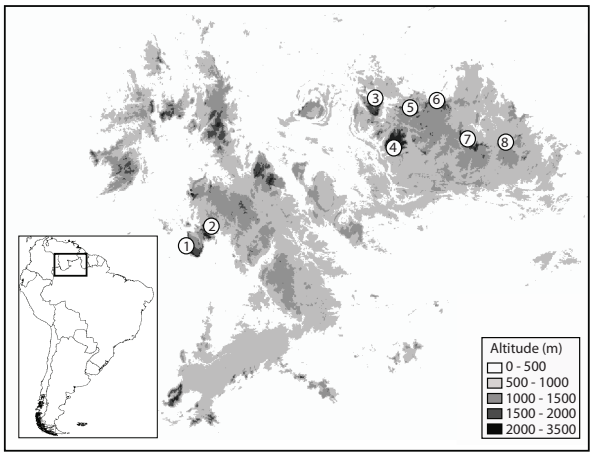

Figure 1.1. Map of the distribution of *Tepuihyla* within the Pantepui of the Guiana Shield: 1. *T. aecii* on Duida tepui; 2. *T. luteolabris* on Marahuaka tepui; 3. *T. edelcae* on Auyan tepui; 4. *T. edelcae* on Chimantá massif; 5. *T. rimarum* on Ptari tepui; 6. *T. rodriguezi* on Sierra de Lema; 7. *T. galani* on Guadacapiapu tepui; 8. *T. talbergae* at Kaieteur falls.

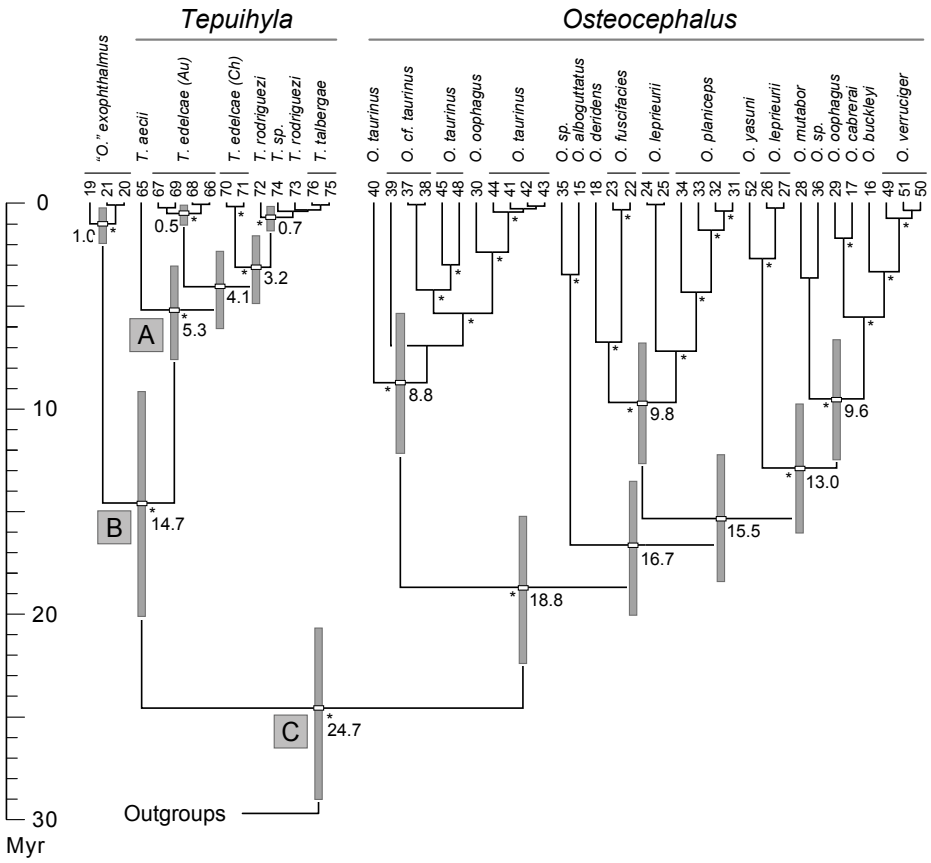

Figure 1.2. Divergence time estimates and 95% confidence bars for *Tepuihyla* (node A), the sister species '*O.' exophthalmus* (node B), and their lowland closest relatives, *Osteocephalus* (node C) for the mtDNA dataset. Divergence estimates of major clades are shown for outgroups. Estimates were obtained using the Lemmon et al. (2007) rate of substitution. Asterisks indicate Bayesian posterior probabilities greater than 95%. The tree and estimates were obtained from BEAST. Numbers next to terminal taxa refer to individual specimens in Table 1.1.

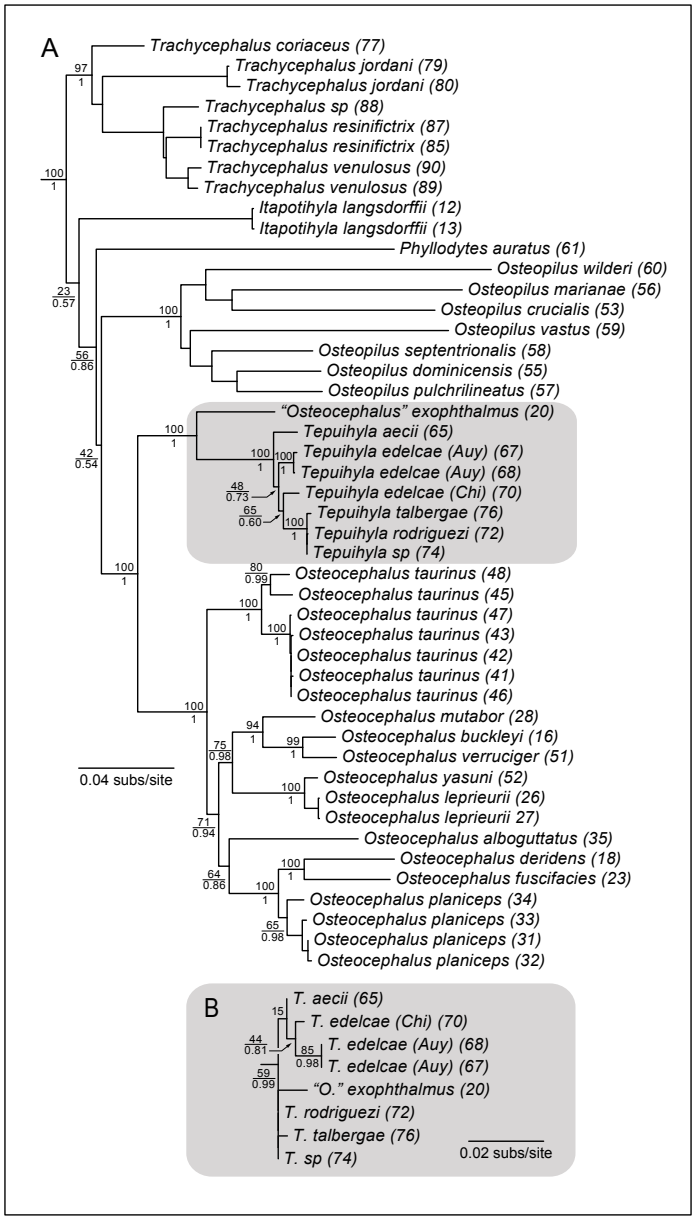

Figure 1.3: Maximum likelihood topology for: (A) the combined dataset of 12S, 16S and POMC (present in all terminals), and (B) the POMC only dataset. Bootstrap values are shown above the nodes and Bayesian posterior probabilities are shown below. Support values are omitted for the outgroups for simplicity. Numbers next to terminal taxa refer to individual specimens in Table 1.1.

Calibration	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
Pseudacris Estimate (0.00277 subs/site/Myr) 12S–16S; 35M gen.	0.7* 0.2–1.4	3.2 1.6–4.9	4.1 2.4–6.2	5.3 3.1–7.7	14.7 9.2–20.2	24.7 20.8–29.2
Xenopus Estimate (0.00249 subs/site/Myr) 12S–16S; 35M gen.	1.1 0.3–2.2	5.0 2.5–8.1	6.4 3.5–9.9	8.1 4.5–12.4	23.6 13.4–34.1	41.5 32.9–51.0
Hylid Fossil Calibrations + GAARlandia 12S–16S; 60M gen.	0.8* 0.2–1.6	3.3 1.5–5.6	4.4 2.2–7.1	5.7 3.0–9.1	17.0 9.6–24.9	31.1 26.5–35.0
Hylid Fossil Calibrations + GAARlandia 12S–16S + POMC; 50M gen.	0.7 1.1–1.4	3.4 1.8–5.4	4.4 2.4–6.6	5.6 3.2–8.5	17.1 10.3–24.3	29.1 23.5–34.2

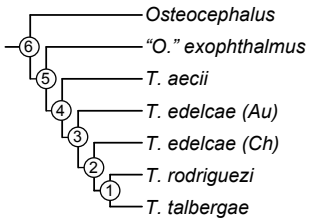

Figure 1.4: Divergence time estimates (top) in millions of years and 95% confidence intervals (bottom) obtained in all BEAST analyses for the main clades of the highland group, *Tepuihyla*, and its closest relatives, *"O." exophthalmus* and *Osteocephalus*. The clades are numbered in the simplified cladogram to the left. Node 1 was not always recovered, and thus we report the deepest divergence within that group and indicate the unrecovered node with *. The differences in number of generations in the left column are due to the fact that different analyses required different numbers of samples to reach stationarity. All divergence time estimates obtained with the calibrations of the outgroup fall within the 95% confidence intervals obtained with the previously estimated rates of evolution.

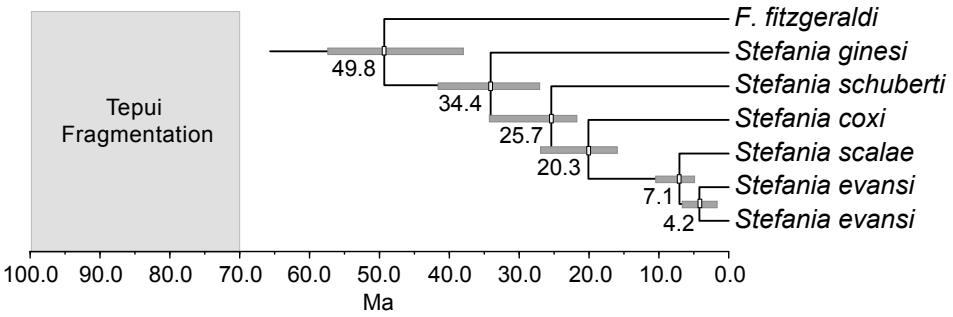

Figure 1.5: Divergence time estimates and 95% confidence intervals for *Stefania*. Divergence time estimates are shown in millions of years. Time of tepui fragmentation is shown as a reference. Estimates were obtained using the Lemmon et al. (2007) rate of substitution.

Chapter 2: Recent evolutionary history of Lost World endemics: population genetics, species delimitation, and phylogeography in a highly complex sky-island landscape

ABSTRACT

The tepuis of South America are massive flattop mountains with cliffs up to 1000m and summits up to 3100m. Tepuis hold enormous endemicity levels, but little is known about the origins of the endemic flora and fauna. Recently diverged lineages offer the possibility of understanding the origins of summit endemicity by examining population dynamics and dispersal. We examine species delimitation, clade relationships, and demographic patterns of three recently diverged lineages of *Tepuihyla*, an endemic treefrog clade. We find unexpectedly high levels of lineage sorting given the low divergences in both nuclear and mitochondrial genes among lineages. We also find overall lower genetic diversity in mitochondrial genes than in nuclear genes in two of the three lineages, which suggests the possibility of a recent mitochondrial selective sweep. We suggest that the genetic and distribution patterns of the four most recently diverged *Tepuihyla* lineages support a concurrent speciation event for these clades. Species delimitation analyses reveal a cryptic, undescribed summit species.

INTRODUCTION

Montane regions harbor many of the global hotspots of diversity, yet the extent of diversity is not fully appreciated because many regions are inaccessible. The flattop mountain area (tepuis) of northern South America make up an endemicity hotspot that

remains largely unexplored. Researchers traditionally believed that the summit biota of the tepuis—the "Lost World" that inspired Conan Doyle's book of the same name—has been isolated atop summits for millions of years. However, molecular analyses indicate that the Lost World is not as isolated and "prehistoric" as popularly thought (Givnish et al., 1997; Kok et al., 2012; Salerno et al., 2012; Bonaccorso and Guayasamin, 2013), though much remains to be done to elucidate the evolutionary history of this fascinating sky-island ecosystem and the causes of its enormous endemicity.

The tepuis are remnants of the Precambrian Guiana Shield plateau, which encompasses a large area of South America east of the Andes and north of the Amazon River basin. These sandstone table mountains were formed approximately 60–90 mya after many cycles of erosion of the Guiana Shield plateau, starting around 300 mya. These remnants form hundreds of sky-islands that reach up to 3000m, with walls up to 1000m high. Hundreds of kilometers of drastically different lowlands separate most summits (Briceño et al., 1990; Briceño and Schubert, 1990; Gibbs and Barron, 1993). Thus, tepuis form a discontinuous ecosystem of sky-islands collectively called Pantepui (Mayr and Phelps, 1955; Huber, 1988), similar to yet arguably of much more extreme topography than other well-known sky-island systems such as the Rocky Mountains and the Western Ghats (DeChaine and Martin, 2005; Smith and Farrell, 2005; Robin et al., 2010).

The Pantepui holds enormous endemicity of many taxa, particularly frogs (~77%; McDiarmid and Donnelly, 2005) and plants (~60%; Huber, 1988; Berry and Riina, 2005). Traditionally, the high endemism has been explained by the Lost World Hypothesis

(Chapman, 1931; Maguire, 1970; Rull, 2004; McDiarmid and Donnelly, 2005). However, many radiations show little to no support for this hypothesis; accumulating evidence strongly supports recent dispersals long after tepuis were formed (Mayr and Phelps, 1967; Givnish et al., 1997, 2011; Rull, 2004; Rull and Nogué, 2007; Salerno et al., 2012; Kok et al., 2012; Bonaccorso and Guayasamin, 2013).

Tepuihyla is a treefrog group of seven endemic Pantepui species. The biogeographic patterns of *Tepuihyla* support the Lost World hypothesis, since all extant species are associated with the Pantepui and its remnants (highlands and midlands) and are mostly allopatric. However, divergence time estimates have revealed *Tepuihyla* to have very recent among-species divergences, supporting dispersal long after tepui formation (Salerno et al., 2012).

Several studies have estimated phylogenetic relationships for members of this group (Kok et al., 2012; Salerno et al., 2012; Jungfer et al., 2013) and briefly suggested that *Tepuihyla edelcae* may consist of two allopatric species that are not sister-groups. However, these analyses either employed non-parametric approaches such as parsimony (Jungfer et al., 2013), had low within-species sampling (Kok et al., 2012; Salerno et al., 2012; Jungfer et al., 2013), or used only a single mtDNA gene (Kok et al., 2012). Support for among-clade relationships was low in all studies, and sampling was not sufficient to test whether the apparent paraphyly of *T. edelcae* was due to sampling error, or if *T. edelcae* indeed represents two evolutionary lineages.

The very recent divergences of *Tepuihyla* likely make the reconstructions of lineage relationships challenging due to factors such as incomplete lineage sorting, which

are not addressed with traditional methods of tree reconstruction. Greater sampling of loci and use of coalescent species-delimitation methods should improve the estimates of relationships among these recently diverged lineages and clarify the evolutionary history and systematics of these Pantepui endemics.

We focus on three recently diverged *Tepuihyla* lineages: *T. edelcae* (from Auyán-tepui), *T. cf. edelcae* (from Chimantá Massif), and *T. rodriguezi*. Using increased sampling of loci and taxa, we performed bayesian concatenated analyses, coalescent gene-tree reconstructions, and species-delimitation analyses to further elucidate species relationships. Population genetic analyses were used to infer recent demographic history and evaluate concordance of patterns across loci within the biogeographic context of the tepui landscape.

MATERIALS AND METHODS

Genetic samples and sequences

The ingroup includes 49 samples within *Tepuihyla*. The samples represent four recognized *Tepuihyla* species: 14 samples of *Tepuihyla edelcae* Auyán from a single locality (southern Auyán-tepui summit), 19 samples of *T. edelcae* Chimantá from three tepui summits on the Chimantá massif (six from Eruoda-tepui, six from Abakapá-tepui, and seven from Churí-tepui, all accessed by helicopter), 12 samples of *T. rodriguezi* from several localities in low and mid-elevations of Venezuela and Guyana, a single *T. aecii* from Cerro Duida, and three *T. exophthalmus* from two lowland localities.

Tepuihyla rodriguezi includes the recently synonymized names *Tepuihyla galani* and *Tepuihyla talbergae*. The species *T. warreni*, which is outside of our focal group, was excluded because of the paucity of GenBank data. The outgroup consists of four species of the sister-lineage *Osteocephalus* (*O. leprieurii, O. taurinus, O. deridens, and O. planiceps*).

We sequenced two mitochondrial segments (12S and 16S ribosomal rDNA genes, 1170bp; ND1, 1160bp) and three nuclear loci (POMC, 480bp; RAG-1, 460bp; Rhodopsin; 835bp). The combined dataset for mitochondrial and nuclear sequences included 4105bp. The GenBank accession numbers for new and previously published sequences are provided in Table 2.3.

The protocols for DNA extraction, PCR amplification (including primers), and sequencing for 12S and POMC are identical to those of Salerno et al. (2012). The thermocycler protocol for RAG-1 and Rhodopsin were the same as for POMC (Salerno et al., 2012), and the protocol for ND1 was the same as for 12S–16S (Salerno et al. 2012). RAG-1 was amplified using the primers from Faivovich et al. (2005), and ND1 was amplified using primers from Moen and Wiens (2009). Rhodopsin was amplified with the primers Rhod1U (5'-AACGGAACAGAAGGCCCAAACTT-3') and Rhod 1L (5'-GCCAAAGCCATGATCCAGGTGA-3'; Pauly 2008).

Gene tree and species tree estimation

We estimated a Bayesian tree with MrBayes v3.2.2 (Ronquist et al., 2012), using 10 million generations, four chains, two replicate runs, and a 10% burnin. We evaluated

convergence and stationarity in Tracer v1.6 (Rambaut and Drummond, 2013). Given that *Tepuihyla* species have very low sequence divergence (Kok et al., 2012; Salerno et al., 2012), the coalescent method implemented in *BEAST (Heled and Drummond, 2010) is appropriate. This coalescent method estimates species trees, taking into account different gene trees as well as population demographic parameters. Unlinked parameters were used for all loci except for the two mitochondrial genes (12S and ND1) that were treated as having linked topologies. We used the MrBayes topology as the input for a priori assignment of clade membership for the *BEAST analysis. Using BEAST 2 (Bouckaert et al., 2014) we performed three independent runs of 120 million generations with a 10% burnin to reach appropriate ESS values (>200). We also evaluated the parameters using TRACER v1.6 (Rambaut and Drummond, 2013) to assess stationarity and convergence of runs. For the *BEAST and the MrBayes analyses, we estimated the best model of evolution and partitioning scheme in PartitionFinder (Lanfear et al., 2012) using the Akaike Information Criterion and treated branch lengths as unlinked. The best partitioning scheme and models of evolution were: 12S (GTR+G), ND1 (GTR+I), POMC 1st+2nd positions (HKY+G), POMC 3rd position (GTR+G), RAG-1 1st+2nd positions (HKY+G), RAG-1 3rd position (K80+I), Rhodopsin Intron (HKY+I), Rhodopsin Exon 1st+2nd position (HKY), Rhodopsin Exon 3rd position (HKY).

We calculated pairwise uncorrected genetic distances (p-distances) with Mega 5.0 (Tamura et al., 2011) using only the four most recently diverged clades (*T. aecii, T. edelcae* Auyán, *T. edelcae* Chimantá, and *T. rodriguezi*).

Species delimitation analyses

Species delimitation analyses were performed using BP&P v2.1 (Rannala and Yang, 2003; Yang and Rannala, 2010), which accommodates the species phylogeny as well as incomplete lineage sorting due to ancestral polymorphism. BP&P requires an input topology of species relationships. Because there is uncertainty of the relationships among the most recently diverged *Tepuihyla* species, and because the species inferences can be extremely sensitive to incorrect input topologies (Leaché and Fujita, 2010), we eliminated *T. aecii*, which is represented by one sample. We input two general hypotheses: (1) *T. edelcae* Chimantá and *T. edelcae* Auyán as sister-groups, and (2) *T. edelcae* as non-monophyletic with *T. rodriguezi* as the sister-group of *T. edelcae* Chimantá (Figure 2.4). *Tepuihyla exophthalmus* is the outgroup in both cases. We also performed seven-taxon analyses and four-taxon analyses (Figure 2.4) to evaluate the effect of *a priori* partitioning of clades, and to examine if the separate summit localities of *T. edelcae* Chimantá and geographically distant populations of *T. rodriguezi* were supported as separate clades. We used both speciation algorithms (0 and 1) with different combinations of priors to confirm stability across runs as suggested by the authors (Yang and Rannala, 2010). Thus, we ran algorithm 0 with three different ε priors (2, 5, 20) and algorithm 1 with four different combinations of α and m (α = 1, 2; m = 0.5, 2). This was done for each input topology. We also performed analyses manipulating the gamma prior G for the population sizes (thetas) to assume either a small (2, 2000) or large (1, 10) ancestral population size. The gamma prior of the age of the root in the species trees (tau) was also manipulated to evaluate the effect of assuming a shallow divergence G (2, 2000)

or deep divergence G (1, 10). The combination of large ancestral population sizes and shallow divergences is assumed to be the most conservative, leading to a lower number of speciation events (Yang and Rannala, 2010). All analyses were performed using a burnin of 500 samples and a total run length of 10,000 samples. We used an ESS value of >200 to determine whether the markov chains reached stationarity.

Population genetics

The population genetic analyses were performed for each of the three putative species: *T. edelcae* Auyán (14 individuals from one summit locality), *T. edelcae* Chimantá (19 individuals from three separate summits atop the Chimantá massif; Figure 2.1), and *T. rodriguezi* (12 individuals of *T. rodriguezi* and the recently synonymized species *T. galani* and *T. talbergae*). *Tepuihyla aecii* was excluded since only one specimen was sampled.

To screen loci for appropriateness as markers, we examined recombination within the loci using the GARD algorithm (Kosakovsky Pond et al., 2006) in DataMonkey (Delport et al., 2010); no loci showed evidence of recombination. To infer demographic history, we estimated number of polymorphic sites, number of haplotypes, nucleotide diversity, and theta using Arlequin 3.5 (Excoffier and Lischer, 2010). We also tested for neutrality of markers using Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) in Arlequin. These commonly used tests assume neutrality, but significant results do not distinguish between non-neutrality and shifts in demographic parameters (Fu, 1997; Excoffier and Lischer 2010).

Nuclear haplotypes were estimated in PHASE 2.1 (Stephens et al., 2001; Stephens and Donnelly, 2003). Because missing data affect the success of haplotype phasing and detection of identical sequences, we reduced all the individual gene alignments to have complete datasets for all loci, which included deleting characters (base pairs) as well as individual samples. No more than six individuals were eliminated from any single gene matrix. Identical haplotypes were eliminated in COLLAPSE 1.2 (Posada, 2004) in order to calculate haplotypes per population for input in Arlequin. We estimated haplotype networks using the minimum spanning network algorithm in Arlequin v3.5, and we used HapStar v0.7 (Teacher and Griffiths, 2011) to edit the network figures.

RESULTS

Phylogenetic and species tree reconstructions

The concatenated MrBayes phylogeny was generally consistent with previous estimates (Figure 2.2). *Tepuihyla rodriguezi*, *T. edelcae* Auyán, and *T. edelcae* Chimantá were each recovered as monophyletic (BPP = 1). For ease of discussion, we refer to these as the three "putative species." Relationships among the three putative species and *T. aecii* are poorly supported (BPP = 0.51 and 0.55). Support for non-monophyly of *T. edelcae* sensu lato (*T. edelcae* Chimantá + *T. rodriguezi*) is low (BPP = 0.55). However, support for monophyly of *T. edelcae* is extremely low, since it was only found in 12.9% of post-burnin trees.

The *BEAST species tree (Figure 2.3A) recovered a sister-group relationship of *T. rodriguezi* and *T. edelcae* Chimantá with moderate support (BPP = 0.76), and a sister-

group relationship for *T. aecii* + *T. edelcae* Auyán with low support (BPP = 0.47). *Tepuihyla exophthalmus* was always recovered (BPP = 1.0) as the sister lineage to all other *Tepuihyla*. All gene tree reconstructions (Figure 2.3B) show that even though some trees support monophyly of *T. edelcae*, this relationship is weak, and that the most common reconstruction of consensus trees (Figure 2.3C) is *T. rodriguezi* + *T. edelcae* Chimantá.

All comparisons between pairs of individuals from the three putative species were 1.5–2.0%. (Table 2.2). The greatest among-clade distance among these three lineages (2.0%) was found for pairwise comparisons between *T. edelcae* Auyán and *T. edelcae* Chimantá. This distance was similar to the distances of *T. aecii* with *T. edelcae* Auyán and Chimantá (1.7–2.0%), and the highest overall distances were found between *T. aecii* and *T. rodriguezi* (2.4–2.5%). All pairwise distances within a putative species were less than 0.5%.

Species delimitation

With the exception of a single set of priors for algorithm 1 (which did not reach appropriate ESS values for most parameters), all analyses for all three input topologies yielded 100% speciation probabilities for *T. edelcae* Auyán, *T. edelcae* Chimantá, and *T. rodriguezi*. These results were found regardless of the input topology of the four-taxon guide tree for the alternative hypotheses: monophyletic *T. edelcae*; (Figure 2.4B) and non-monophyletic *T. edelcae* (Figure 2.4C).

Speciation within *T. edelcae* Chimantá (among Eruoda, Abakapá, and Churí tepuis) and within *T. rodriguezi* (Venezuela and Guyana) was evaluated using the seven-taxon species tree (Figure 2.4A). The speciation probabilities within Chimantá varied substantially with changing priors, particularly when changing the theta prior from G (1, 10) to G (2, 2000). As expected, the most conservative prior combination for theta G (1, 10) and tau G (2, 2000) resulted in the lowest speciation probabilities within Chimantá, regardless of speciation algorithm and priors, and these speciation probabilities were all between 0.37–0.50. Speciation probabilities for *T. rodriguezi* were always moderate to very high (0.88–1.00).

Population genetics and demographics

The Chimantá population (all three tepuis combined) had the smallest number of polymorphic sites and number of haplotypes among the three nuclear loci (Table 2.1). The two mitochondrial genes had similar numbers of haplotypes and polymorphic sites for Chimantá and Auyán; *T. rodriguezi* had the fewest 12S haplotypes. All haplotypes found for each of the three putative species were unique to that population with the exception of one haplotype of RAG-1, which was the most common haplotype for all three populations. None of the individual tepuis atop the Chimantá massif had unique haplotypes when analyzed separately. When *T. rodriguezi* was analyzed as two populations (Venezuela and Guyana), some unique haplotypes were found in each locus.

A high degree of lineage sorting was found in the nuclear loci (Figure 2.5). All but one haplotype are unique to each putative species. Rhodopsin is completely sorted

among the three putative species. POMC is almost completely sorted; some POMC haplotypes of Auyán are more closely related to Chimantá haplotypes than to other Auyán haplotypes. That is, the haplotype tree of Auyán is paraphyletic to that of Chimantá (Figure 2.5D). RAG-1 has an unsorted haplotype (Figure 2.5C) that is shared in equal proportions among all three groups; all other RAG-1 haplotypes are uniquely derived in each putative species, many directly derived (single mutation) from that common shared haplotype.

For each putative species, the range of mtDNA nucleotide diversity (12S and ND1) is lower than, or overlaps with, that for the three nDNA genes. For *T. edelcae* Auyán and *T. rodriguezi*, the number of mtDNA haplotypes/sample overlaps with that for the three nDNA genes. In Chimantá, however, the number of haplotypes/sample is as least twice as much for mtDNA than for nDNA (Table 2.1).

Almost all test statistics were negative, but many were not statistically significant. Fu's Fs statistic showed several highly significant values, but the distribution of these across loci and populations showed no general pattern. Only two Tajima's D values were significant (12S; Table 2.1); Rhodopsin showed no significant results for Tajima's D or Fu's Fs.

DISCUSSION

Phylogenetic reconstructions

The concatenated MrBayes tree was concordant in topology and degree of support with published phylogenies (Salerno et al., 2012; Kok et al., 2012; Jungfer et al., 2013).

We found very strong support (BPP = 1) for the existence of three monophyletic lineages (Figure 2.2): *T. edelcae* Chimantá, *T. edelcae* Auyán, and *T. rodriguezi*. The support for non-monophyletic *T. edelcae* (*T. edelcae* Chimantá + *T. rodriguezi*) is moderately low (BPP = 0.55), but the support for monophyletic *T. edelcae* is extremely low (BPP = 0.13). The *BEAST coalescent species reconstruction method, which is perhaps more appropriate given the shallow divergences, yielded results very similar those of the MrBayes tree (Figures 2.2, 2.3A). The DensiTree visualization and the consensus trees from *BEAST (Figure 2.3B and C) suggest that the problematic reconstructions are due to the placement of *T. aecii* and *T. edelcae* Auyán. Even though the latter relationships are unresolved, the reconstruction of *T. rodriguezi* + *T. edelcae* Chimantá is the most commonly obtained, supporting the hypothesis that *T. edelcae* is paraphyletic regardless of other relationships in the group.

Increased sampling of loci and taxa did not significantly improve phylogenetic estimation of among-species relationships, even when using coalescent species-tree inference (*BEAST). It is possible that if these lineages diverged during near-simultaneous speciation events, then more genetic data will not resolve relationships among these (Karl et al., 2012).

Unexpected pattern of nuclear vs. mitochondrial haplotype diversity: possible interpretations

Mitochondrial genes should in general coalesce four times faster than nuclear genes (Avise, 2000; Palumbi et al., 2001). Thus, it is expected that low mitochondrial

divergence among lineages be accompanied by a high degree of incomplete lineage sorting in the nuclear genome (Avise, 2000). The uncorrected pairwise genetic distances among individuals of the three putative species of *Tepuihyla* are all between 1.5–2.0% for mtDNA (ND1 and 12S combined), which is generally considered low for between-species comparisons (Vences et al., 2005; Fouquet et al., 2007). In contrast, the high levels of lineage sorting and unique nuclear gene haplotypes indicate that in spite of the low divergence in mtDNA, these lineages have been evolving largely independently.

mtDNA has long been considered a neutral or nearly neutral marker ideal for reconstructing recent demographic history due to its fast rate of evolution (in animals), its much smaller effective population size, and apparent lack of recombination (Ballard and Kreitman, 1995; Avise, 2009). However, many studies have suggested (and sometimes strongly advocated) that mtDNA alone is not a neutral marker (Ballard and Whitlock, 2004; Bazin et al., 2006; Leaché and McGuire, 2006; Meikeljohn et al., 2007; Dowling et al., 2008; Galtier et al., 2009; Rato et al., 2010; Messer and Petrov, 2013), and in fact may be under recurrent selective pressures, particularly for species with very large effective population sizes (Bazin et al. 2006).

Under normal assumptions (constant population sizes, no incomplete lineage sorting, no introgression, and no selection) the expectation is that mitochondrial loci should have greater diversity than nuclear genes (in vertebrates). The haplotype networks (Figure 2.5) and population parameters (Table 2.1) either show no evidence for this, or show the opposite.

Three scenarios might explain the unexpectedly low mtDNA haplotype diversity: recent population expansion, male-biased migration, and selective sweeps (acting either on mtDNA or nDNA). We tested the first scenario, recent population expansion, using Tajima's D and Fu's Fs. The Fs tests found seven significant results whereas D detected only two. This is not unexpected because Fs is more sensitive than D to demographic changes (Fu, 1997; Excoffier and Lischer, 2010; Raposo Do Amaral et al., 2013). There was general agreement in the magnitude and direction of the test statistics for each pair of D and Fs comparisons. However, a few results were inconsistent across loci and populations for the two tests. For example, in *T. rodriguezi* Fs was highly significant (-3.095) for ND1, but not significant (0.297) for 12S. The interpretation of D and Fs is not straightforward. The assumption is that loci and/or polymorphic sites are neutral, since in non-neutral loci significant negative values for both tests may indicate either recent demographic change (population expansion and/or bottleneck) or a selective sweep (Ballard and Whitlock, 2004). The first scenario (demographic change) is not supported with our dataset, particularly given the much higher diversity in nDNA (compared to mtDNA) in two of the three lineages. It is also not fully rejected, particularly in *T. edelcae* Chimantá, in which even though the neutrality test results are not consistent across loci, the genetic diversity is consistently low; thus this is the only lineage where the diversity incongruence is not observed between mtDNA and nDNA.

The second scenario is male-biased migration. Under this scenario male migrants introduce genetic variation to the populations, but only nDNA increases in diversity given that mitochondria are matrilineally inherited. Under this scenario, there must have

been an initial isolation of the populations followed by a reconnection of the lineages where male-biased migration was favored. There are some documented cases of male-biased (Lampert et al., 2003; Palo et al., 2004) and female-biased (Austin et al., 2013) dispersal in anurans inferred from molecular data. These cases are mostly from species where much is known about its genetic and ecological aspects. Thus, in species with little to no information about their ecology and dispersal abilities, this scenario cannot be assessed.

The third scenario is a selective sweep, either through direct selection of mtDNA, or though selection on nDNA resulting in mtDNA hitchhiking through cytonuclear interactions (Ballard and Whitlock, 2004; Dowling et al., 2008). Since we did not sample whole genomes, we cannot directly test this. However, we can weigh the likelihood of these scenarios.

It has been proposed that even low levels of selection on mitochondria may cause hitchhiking of the mtDNA genome (and thus a selective sweep) due to the non-recombining and gene-dense nature of the mitochondrion. These selective sweeps would ultimately result in a reduction of mitochondrial genetic diversity, but not nDNA genetic diversity (Bazin et al. 2006). This scenario is somewhat consistent with our results (Table 2.1). The selective sweep was likely soft rather than hard (Messer and Petrov, 2013), because the large number of missing mitochondrial haplotypes between the populations suggests that the sweep happened after the lineages became isolated and had accumulated mtDNA variation. The sweeps would have reduced standing genetic variation in each population independently, resulting in fixation of different mtDNA haplotypes for each

putative species. Thus, the observed mtDNA haplotype diversity is only what has been accumulated since the selective sweep.

Hence, two hypotheses may explain the patterns of incongruence in mtDNA and nDNA diversity of *T. edelcae* Auyán and *T. rodriguezi*. The first possibility is male-biased migration following population divergence, resulting in an incongruence in diversity between the two genomes. The second possibility is mitochondrial selective sweeps, which would have reduced the genetic diversity in mtDNA relative to nDNA. These hypotheses cannot be distinguished with our dataset or with the sparse ecological data for Pantepui anurans.

Species delimitation and systematic implications

All recognized *Tepuihyla* species were described based on a combination of phenotypic and morphometric differences (Ayarzagüena et al., 1992). However, these are subtle; most highland species are of similar size, dorsum coloration (dark brown/grey), webbing, and general appearance, and differ only slightly from the low- and mid-elevation species, in which the most obvious difference is coloration patterns (Ayarzagüena et al., 1992). No morphological differences have been reported between the two putative species of *T. edelcae* , which were described as a single species with a disjunct distribution, atop Auyán-tepui and many tepuis atop the chimantá massif (Ayarzagüena et al., 1992).

Recent phylogenetic analyses led to the addition of two species to *Tepuihyla*, *T. exophthalmus* and *T. warreni* (Salerno et al. 2012; Jungfer et al., 2013). These two

species are the most deeply diverged within *Tepuihyla*, with the divergence of *T. exophthalmus* estimated to be 15mya (Salerno et al., 2012); both are found in the mid-elevations of the eastern Pantepui (Figure 2.1). The most recent divergences in *Tepuihyla* represent four lineages: *T. rodriguezi* (mid-elevations of the eastern Pantepui), *T. aecii* (highlands of Cerro Duida in western Pantepui), *T. edelcae* Auyán (summit of Auyán-tepui in eastern Pantepui), and *T. edelcae* Chimantá (summits of several tepuis atop the Chimantá massif in eastern Pantepui). Our analyses found extremely low support for *T. edelcae* as a monophyletic group (12.9%), but among-clade relationships are overall poorly supported.

In the seven-taxon species delimitation analysis, the populations atop the three tepuis (Eruoda, Abakapá, and Churí) within the Chimantá massif were not supported as separate evolutionary lineages, particularly when using the conservative speciation priors, where speciation probabilities were all extremely low (0.37–0.5). This is consistent with the concatenated MrBayes tree (Figure 2.2).

There is fairly high support for two lineages within the species *T. rodriguez*i, corresponding to two general localities of Venezuela and Guyana, even when using the conservative combination of priors (the support varies from 0.88–1.0). Frogs at each locality were considered to be distinct species (*T. rodriguezi* in Venezuela and *T. talbergae* in Guyana) until they were synonymized recently (Jungfer et al., 2013). Our molecular data suggest that the synonymy should be re-considered. More extensive sampling and consideration of other lines of evidence (i.e., integrative taxonomic approaches) are needed.

Regardless of the input topology, the Auyán and Chimantá lineages of *Tepuihyla edelcae* are supported as distinct (Figure 2.4A,B). Given the body of evidence, such as monophyly of individual lineages, lack of support for the monophyly of *T. edelcae* Chimantá + *T. edelcae* Auyán, distribution of unique haplotypes, and highly supported speciation probabilities from the BP&P analysis, we argue that *T. edelcae* be considered two distinct species; this question is under further investigation (Salerno et al., *in progress*).

Tepuis as islands is the sky: the case of Tepuihyla

The tepuis and many other fragmented highland ecosystems have been described as islands in the sky, where ecological and climatic conditions are so different that populations on a summit are effectively isolated from other populations. In the case of tepuis, the vertical topography of many of them is arguably unsurpassed in steepness by other mountainous systems, suggesting that tepui species are cut off from other summits and have been for millions of years. Without doubt this has fostered the romanticization of the Lost World as a cradle of biodiversity. However, levels of genetic divergence on distinct tepuis have been shown to be surprisingly low among several species (Salerno et al., 2012; Kok et al., 2012; Bonaccorso and Guayasamin, 2013) indicating that populations have moved between summits long after tepui formation.

Assuming tepuis are sky-islands, whether topographical or ecological, one can make predictions about patterns of genetic diversity. For example, genetic diversity within and among populations should follow general island biogeography and population

genetic predictions: smaller islands should hold smaller effective population sizes and will have lower genetic diversity due to genetic drift and higher fixation rates. Larger distances among islands should result in higher levels of divergence and a higher number of unique haplotypes.

Ecological distances or differences between these sky-islands may function similarly to geographic distances. For example, for species restricted to islands with elevations >2000 m it is likely that the lower elevations are unsuitable or at the very least not preferred, thus increasing the effective distance between these populations. On the other hand, mid- and low-elevation species should have a higher tolerance for lowland conditions, thus reducing the ecological distance between them, even if suitable habitat is patchy.

Given their very low levels of genetic divergence, it is clear that *Tepuihyla* species have overcome topographic barriers such as cliffs and dispersed across the highlands of the Pantepui. Thus it is likely that the cessation of gene flow among the three putative species is due to separation of these by unsuitable habitat and not extreme topography. Climatic and ecological conditions are therefore likely the main forces promoting divergence in a manner similar to other sky-island ecosystems (DeChaine and Martin, 2005; Smith and Farrell, 2005; Robin et al., 2010).

Auyán-tepui and the Chimantá massif, which are part of the Eastern Pantepui province, are separated by about 60km of lowland savannas. An important difference between Auyán and Chimantá is that the latter is a massif supporting ten tepuis. The three tepuis atop Chimantá are 15–30km apart and have extreme intervening landscape features.

Rather than dry lowland savannas, mid-elevation wet tropical forest separates the classic summit Pantepui habitat common in most sandstone tepuis of the Eastern District (Huber, 2006). Although summit habitats, particularly those with Brocchinia peat bog habitats, seem to be preferred by *Tepuihyla* (Ayarzagüena et al., 1992), the intervening wet tropical forest of the Chimantá tepuis is almost certainly more suitable for dispersal of *Tepuihyla* than are the dry tropical savannas separating Auyán from Chimantá. Thus, the topography and habitat suggest that these forests may permit some gene flow among the Chimantá tepuis.

Because only a single locality on Auyán-tepui was sampled, we expected the three Chimantá samples would in sum have much higher haplotype diversity than Auyán. However, we observed the opposite; the single Auyán population has 2–5 times as many nuclear haplotypes as Chimantá has in its three combined populations (POMC: 11 vs 2, RAG-1: 6 vs 4, and Rhodopsin: 4 vs 2). Another expectation is some level of population structure among the Chimantá tepuis, though each population with lower diversity than Auyán (since the summit areas are individually smaller than Auyán). However, none of the three summit populations atop Chimantá has unique haplotypes for any gene. The most likely explanation for the low diversity and lack of structure across the Chimantá summits is that a recent bottleneck decimated the populations (and thus the haplotype diversity), and they only recently started expanding atop the massif. This is consistent with the general pattern of low haplotype diversity (both nuclear and mitochondrial) of Chimantá and the lack of structure among populations, but not consistent with neutrality tests, since only 12S and POMC have significant Fs statistics. Also, this bottleneck

scenario may explain the suggestion that Chimantá has an overall depauperate herpetofauna (McDiarmid and Donnelly, 2005; Señaris and MacCulloch, 2006). The smaller individual summits and the much more complex landscape of Chimantá may in general support much smaller effective population sizes, making Chimantá more prone to stochastic effects and local extinctions compared to Auyán.

One of the most interesting results concerns *Tepuihyla aecii* from Cerro Duida, for which we have only one sample. This massif is located 400 km southwest of Auyán and Chimantá (Figure 2.1), yet its genetic distances to other *Tepuihyla* completely overlap in range with those between *T. edelcae* Auyán and *T. edelcae* Chimantá, even though the geographic distance is almost seven times greater. For example, the distances between *T. aecii* and *T. edelcae* Auyán are 1.7–2.0%. The single *T. aecii* datum suggests a recent, pervasive dispersal of *Tepuihyla* species across the entire Eastern Pantepui region, followed by fragmentation into isolated species.

CONCLUSIONS

The lack of resolution of interspecific relationships among the four clades (*T. aecii*, *T. rodriguezi*, *T. edelcae* Auyán and Chimantá) can be used to generate a hypothesis for the evolution of these lineages. If the relationships are unresolved due to concurrent speciation events, then we can hypothesize that the common ancestor to all four may have been a widespread Pantepui lowland dweller, widespread in both eastern and western Pantepui (and potentially also altitudinally widespread from lowland to summits). At some recent time low- to mid-elevation conditions (whether biotic or

abiotic) may have become unsuitable, at which point most lowland populations became extinct, resulting in concurrent speciation events due to isolation of populations on the tepui summits within the same timeframe. This idea is consistent with our analyses; for example, sequence divergences are very similar among the four lineages regardless of the geographic distance separating them. This scenario is also consistent with studies of pollen deposit data, which suggest recent downward vertical migration of the tepui flora diring glacial maxima in the Pleistocene, allowing for lowland interchange and dispersal to other highlands (Rull, 2004; Rull and Nogué, 2007). The hypothesis of recent concurrent *Tepuihyla* speciation in most of the Pantepui, can serve as a starting point for future studies regarding the origin and evolution of Pantepui taxa.

The case of *Tepuihyla* offers some insights on the recent evolutionary history atop the Lost World. For example, even though the three putative species are very recently diverged and colonized summits long after tepui formation, they are currently evolving independently, as evidenced by almost complete lineage sorting and unique haplotypes for all species and genes. Thus, the complex topography alone does not seem to be the main barrier to gene flow. Local tepui conditions and extreme biotic and abiotic differences among highlands and lowlands, effectively making them an ecological archipelago of islands in the sky, may promote high endemicity atop the tepui summits. Species distribution modeling, least-cost path analyses, whole-genome approaches, and physiological tolerance experiments may be able to elucidate the roles of selection, ecology, and physiology in generating and maintaining the enormous endemicity found in the Pantepui.

Table 2.1. Summary statistics of population parameters obtained in Arlequin for five loci and three populations: *T. edelcae* Auyán, *T. edelcae* Chimantá, and *T. rodriguezi*.

Locus	Population	Sample size	# Polymorphic sites	# Haplotypes	# Haplotypes/ sample size	Nucleotide diversity (π) ± SD	Theta (S)	Tajima's D	Fu's Fs
12S mtDNA	*T. edelcae* Auyán	15	3	4	0.267	0.0006±0.00066	0.923	-1.685*	-2.369**
	T. edelcae Chimantá	19	2	4	0.211	0.0004±0.00062	0.572	-1.511*	-3.670**
	T. rodriguezi	12	1	2	0.166	0.0005±0.00057	0.331	-0.194	0.297
ND1 mtDNA	*T. edelcae* Auyán	13	2	3	0.231	0.0011±0.00141	0.644	-1.468	-1.402*
	T. edelcae Chimantá	11	2	3	0.273	0.0027±0.00242	0.682	0.199	-0.019
	T. rodriguezi	7	2	3	0.429	0.0026±0.00242	0.816	-0.275	-0.438
POMC nDNA	*T. edelcae* Auyán	28	5	11	0.393	0.0047±0.00304	1.285	1.877	-4.120**
	T. edelcae Chimantá	32	1	2	0.063	0.0001±0.00035	0.248	-1.142	-1.265*
	T. rodriguezi	24	5	10	0.417	0.0037±0.00262	1.339	0.411	-5.259**
RAG1 nDNA	*T. edelcae* Auyán	28	4	6	0.214	0.0028±0.00227	1.028	-0.226	-1.934
	T. edelcae Chimantá	32	3	4	0.125	0.0022±0.00191	0.745	-0.241	-0.499
	T. rodriguezi	24	6	9	0.375	0.0044±0.00311	1.607	-0.215	-4.024**
Rhod nDNA	*T. edelcae* Auyán	28	4	4	0.143	0.0021±0.00148	1.028	1.067	1.312
	T. edelcae Chimantá	28	1	2	0.071	0.0002±0.00034	0.257	-0.740	-0.380
	T. rodriguezi	22	5	6	0.273	0.0017±0.00129	1.372	-0.393	-1.581

Table 2.2. Uncorrected pairwise distances within and among putative species, *T. edelcae* Auyán, *T. edelcae* Chimantá, *T. rodriguezi*, and *T. aecii*.

	T. edelcae Auyán	*T. edelcae* Chimantá	*T. rodriguezi*
T. edelcae Auyán (n = 14)	0.001	-	-
T. edelcae Chimantá (n = 19)	0.016–0.020	0.001–0.004	-
T. rodriguezi (n = 12)	0.015–0.017	0.015–0.019	0.000–0.002
T. aecii (n = 1)	0.017–0.020	0.020	0.024–0.025

Table 2.3: Genetic samples and localities. MHNLS = Museo de Historia Natural La Salle, BPN = Brice P. Noonan Field series, PS = Patricia Salerno field series, TNHC = Texas Natural History Collection Field Series.

Species	Field/ museum code	Locality	Coordinates
T. aecii	MHNLS12013	Cerro Duida, Amazonas, Vzla	3°18'N 65°37'W
T. exophthalmus	BPN166	Guyana	5°37.30'N 60°14.42'W
T. exophthalmus	PS205	Sierra de Lema, Canaima, Vzla	5°54.045'N 61°26.290'W
T. exophthalmus	PS206	Sierra de Lema, Canaima, Vzla	5°54.045'N 61°26.290'W
T. edelcae	MHNLS16090	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	PS002	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05824	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05825	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05826	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05827	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05828	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05829	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05830	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05831	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05833	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05834	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05835	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. edelcae	TNHC05836	Auyán-tepui, Canaima, Vzla	5°46.599'N 62°32.251'W
T. cf. "edelcae"	PS268	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS330	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS331	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS332	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS334	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS337	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS339	Churí-tepui, Chimantá massif, Canaima, Vzla	5°15.257'N 62°00.472'W
T. cf. "edelcae"	PS365	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W
T. cf. "edelcae"	PS366	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W
T. cf. "edelcae"	PS367	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W
T. cf. "edelcae"	PS368	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W
T. cf. "edelcae"	PS398	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W

Table 2.3 *continued*

T. cf. "edelcae"	PS399	Abakapá-tepui, Chimantá massif, Canaima, Vzla	5°11.497'N 62°18.939'W
T. cf. "edelcae"	PS410	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. cf. "edelcae"	PS446	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. cf. "edelcae"	PS447	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. cf. "edelcae"	PS448	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. cf. "edelcae"	PS449	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. cf. "edelcae"	PS450	Eruoda-tepui, Chimantá massif, Canaima, Vzla	5°22.525'N 62°05.674'W
T. rodriguezi	BPN1101	Mazaruni-Potaro, Guyana	5°43.389'N 60°16.087'W
T. rodriguezi	BPN1218	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1219	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1220	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1221	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1222	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1223	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1224	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1225	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	BPN1226	Mazaruni-Potaro, Guyana	5°72'N 60°27'W
T. rodriguezi	PS003	Luepa, Bolívar, VENEZUELA	5°44.46'N 61°31.02'W
T. rodriguezi	PS197	Sierra de Lema, Canaima, Vzla	5°49.228'N 61°25.473'W
O. taurinus	PS004	Las Claritas, Bolívar, Venezuela	6°10.49'N 61°25.17'W
O. leprieurii	MHNLS18689	Uey River, Bolívar State, Venezuela	-
O. planiceps	QCAZ19195	Estación Científica Yasuní, Orellana, Ecuador	-
O. deridens	QCAZ20868	Ecuador	-

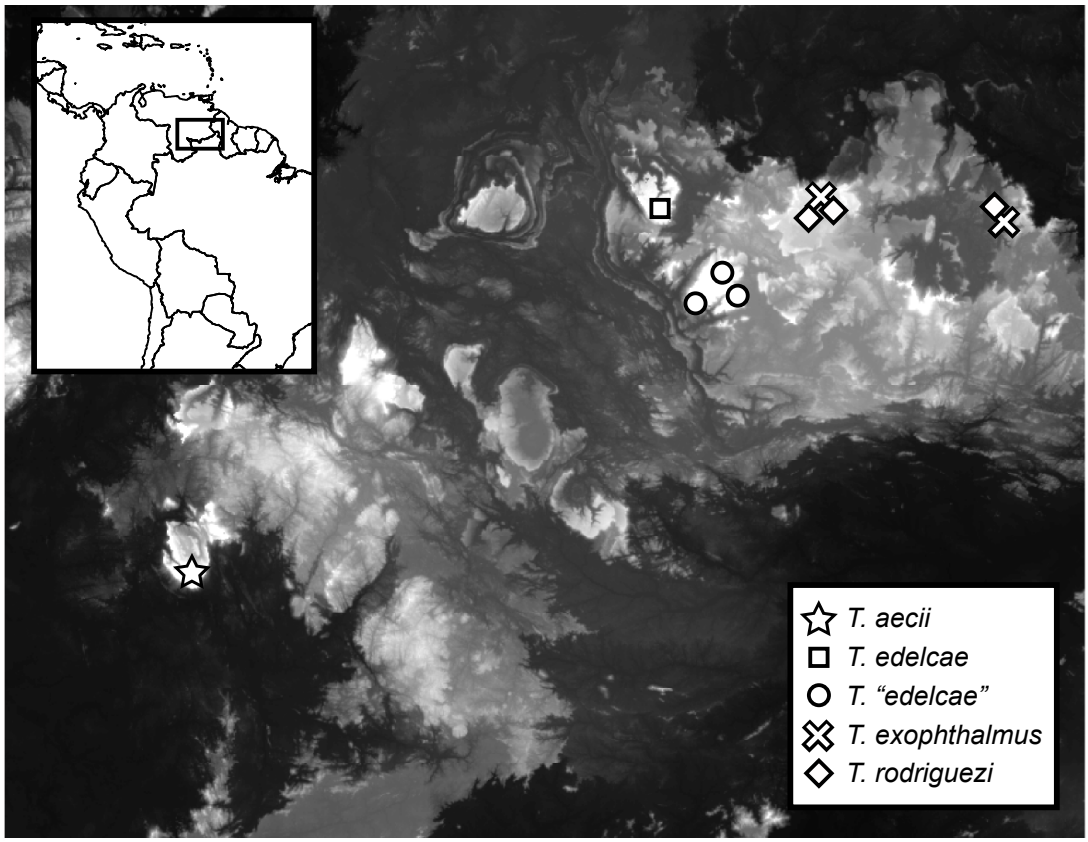

Figure 2.1. Map of eastern Venezuela, showing locations of genetic samples of *Tepuihyla*.

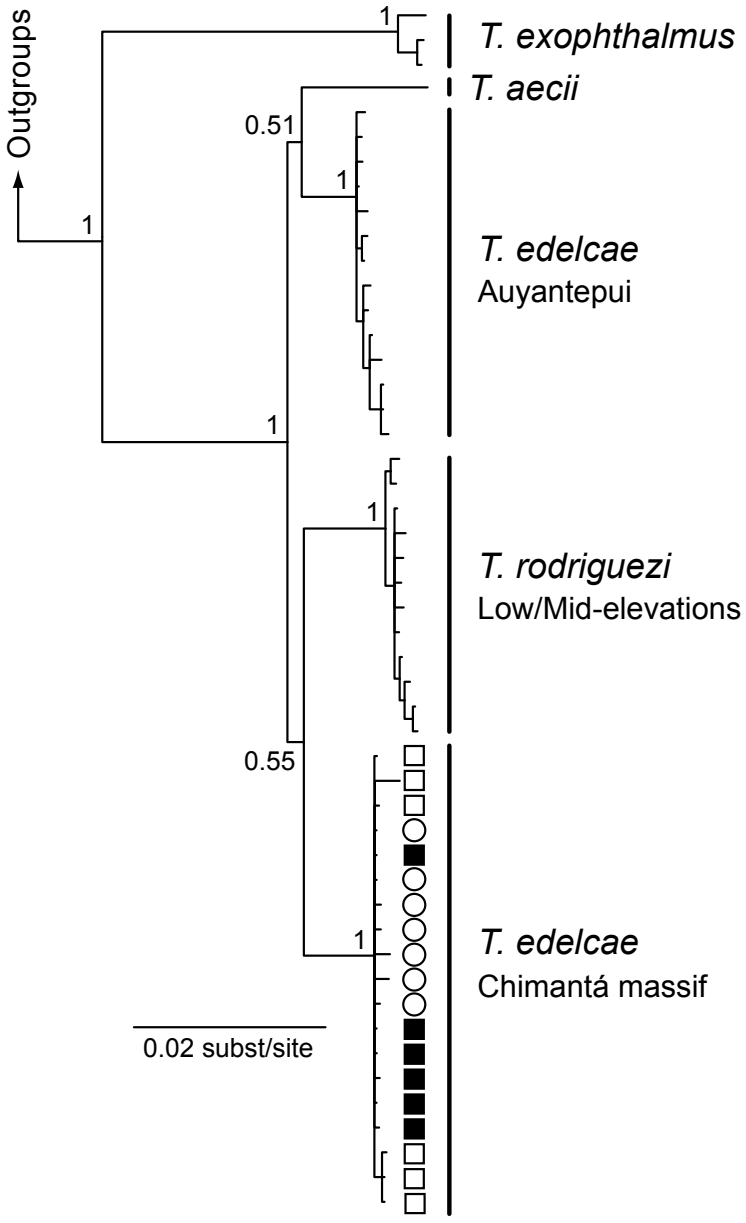

Figure 2.2. Phylogenetic reconstruction of the concatenated dataset obtained in MrBayes. Bayesian posterior probabilities are shown at nodes. Outgroups not shown.

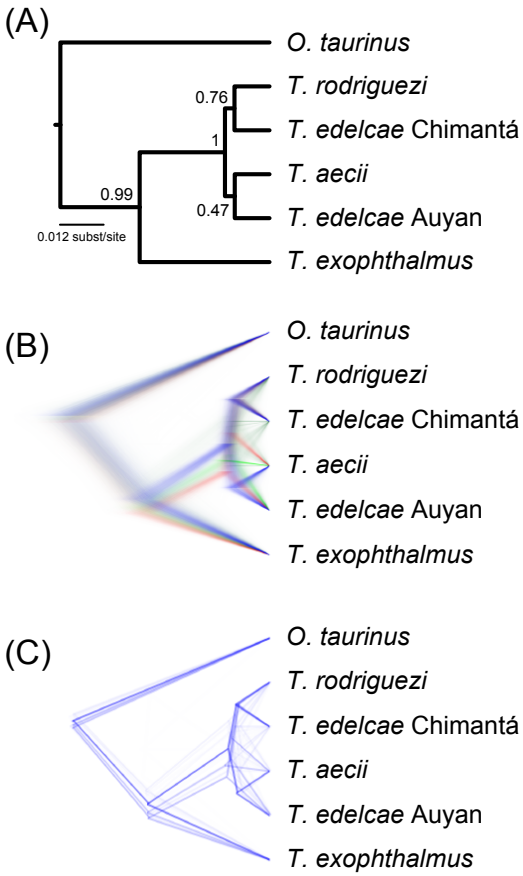

Figure 2.3. Coalescent species reconstruction in *BEAST. (A) Species tree with posterior probabilities, (B) DensiTree visualization of all estimated gene trees, and (C) DensiTree visualization of possible consensus trees for the given loci.

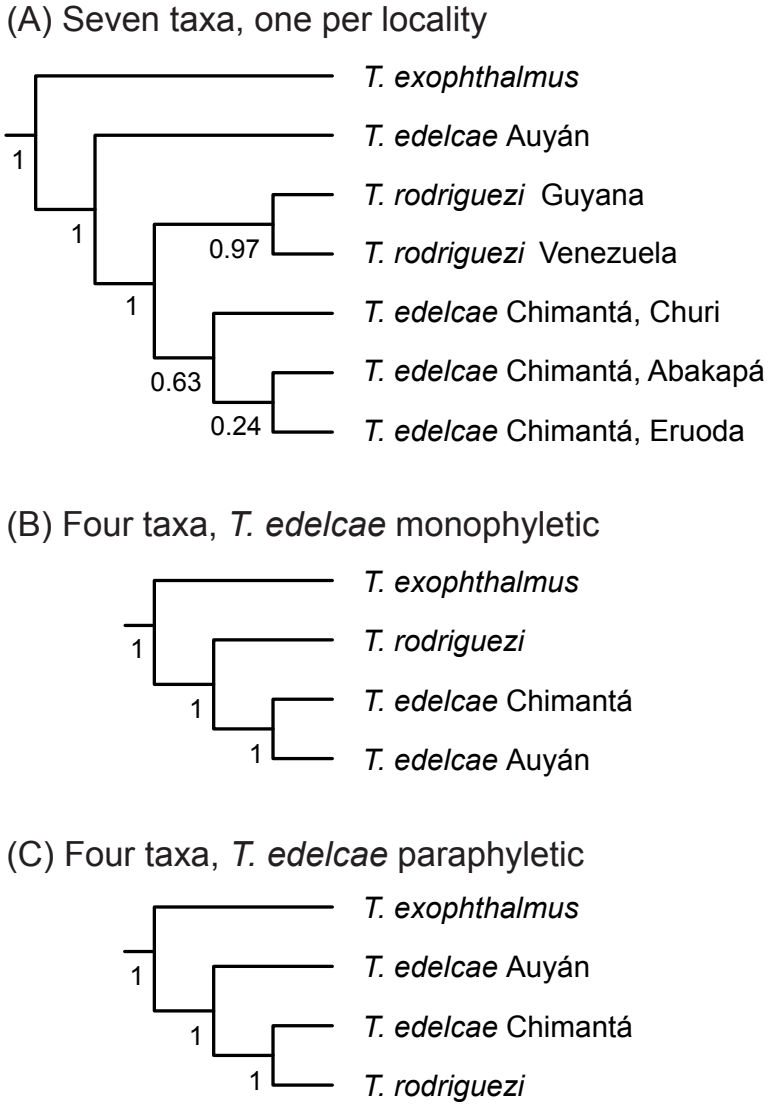

Figure 2.4. Species delimitation analyses performed in BP&P. (A) Reconstruction given the seven-taxon input tree, and (B) and (C) are the four-taxon input tree, with the two alternative hypotheses (*Tepuihyla edelcae* as monophyletic and *T. edelcae* as non-monophyletic, respectively).

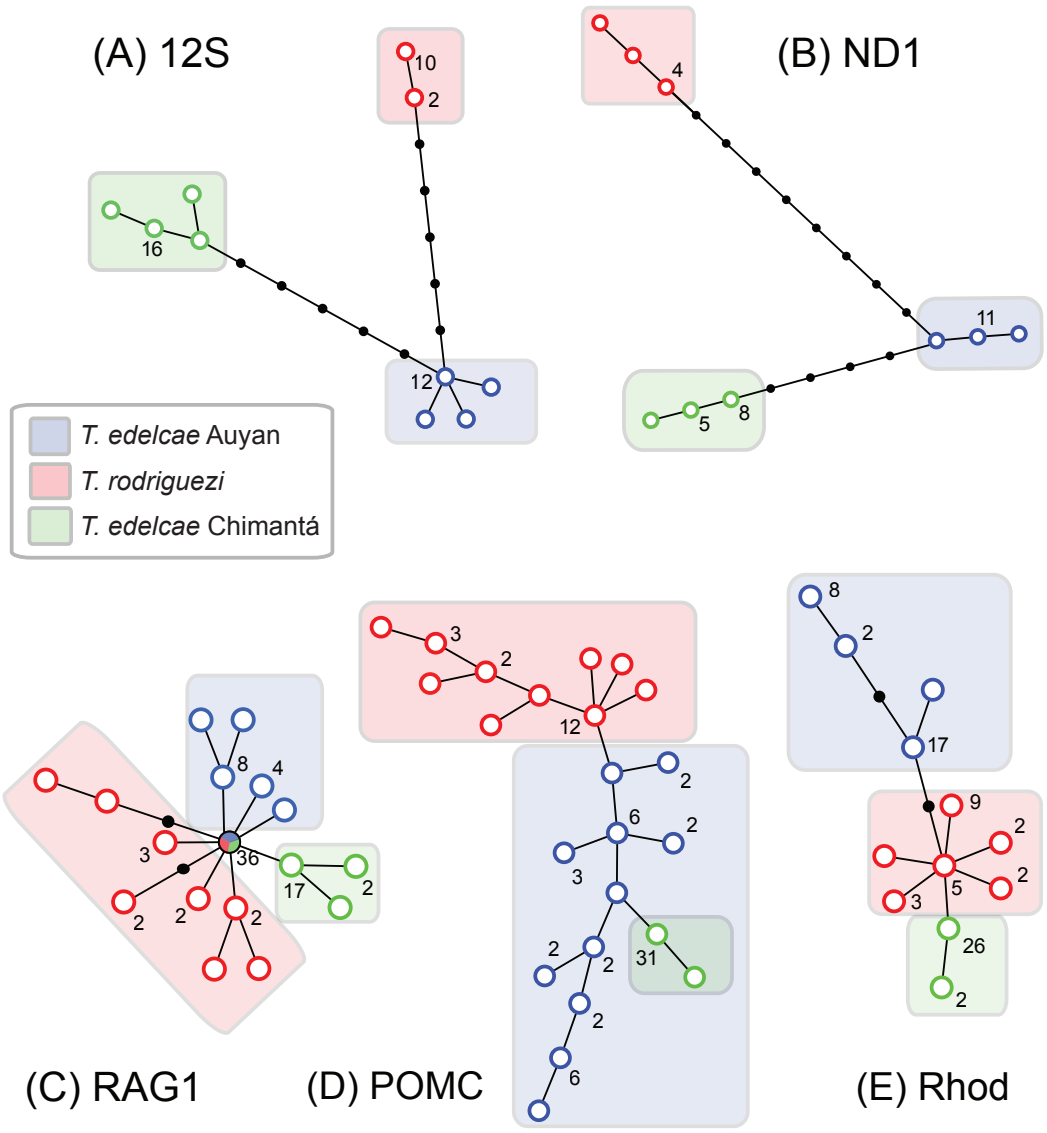

Figure 2.5. Haplotype networks obtained in Arlequin for all five loci: (A) 12S mitochondrial rRNA, (B) ND1, (C) RAG-1, (D) POMC, and (E) Rhodopsin. Black dots represent missing haplotypes and tic marks represent mutations. Colored circles represent haplotypes in populations; the number of individuals with that haplotype is represented as a number next to the circle. Colors identify the three different populations.

Chapter 3: Population genomics of an endemic Lost World summit frog

ABSTRACT

The Lost World of South America, made up of hundreds of massive flattop mountains, has some of the most drastic landscapes on earth, and holds some of the highest levels of endemicity in groups such as frogs and plants. The endemic frog, *Tepuihyla edelcae*, is restricted to the highlands of these flattop mountains, and is a recent colonizer of the tepui summits. Here, we use next-generation sequencing technologies to examine population structure atop these formations, and thus determine effects of landscape on connectivity and isolation. Using STRUCTURE and Discriminant Analysis of Principal Components, we find high levels of population assignment and structure in spite of low mtDNA divergences. Given previous data on the species, our data suggest that there has been rapid accumulation of variation unique to each tepui summit in spite of a very recent bottleneck and population expansion that took place in these formations.

INTRODUCTION

Most hotspots of endemism and biodiversity are located in islands or mountainous systems (Orme et al. 2005). High-elevation mountain systems are thus often referred to as sky islands, since their biotic and abiotic conditions present unique challenges that effectively make them ecological islands (Mayr and Diamond, 1976; Masta, 2000; Knowles, 2001; DeChaine and Martin, 2005; Bech et al., 2009; Robin et al., 2010). The Lost World of South America has arguably one of the most drastic landscapes on earth.

With sheer cliffs of up to 1000m, summit elevations of up to 3100m, and lowland separations of sometimes hundreds of kilometers between them, the tepuis (flattop mountains of northern South America) form a complex sky-island archipelago (Huber, 1988; 2006). The ecological differences between summits and lowlands are often as extreme as the landscape, with lowland dry savannas and tropical forests in the foothills, while the summit biota is characterized mostly by peat bogs and small shrubs (Mayr and Phelps, 1967; Huber et al., 2001; Huber, 2006).

Although the summits have nutrient-poor soils, drastic daily temperature fluctuations, and no permanent bodies of water (Huber, 1988; 1992; 1995; 2006; Huber et al., 2001), the Lost World holds some of the world's highest levels of endemism for groups such as frogs and plants (Berry and Riina, 2005; McDiarmid and Donnelly, 2005). The reasons for the high endemism, however, are not well established, and the relative contributions of topography and ecology to summit isolation are not clear.

Understanding the effect of landscape on gene flow and population structure is a key issue in phylogeography, speciation, conservation genetics, and population genetics. The effects of naturally occurring geographic barriers on population structure have been long studied, but still remain highly debated topics, particularly when attempting to understand their role in generating and maintaining biodiversity (Moritz et al., 2000; Jetz et al., 2004; Orme et al., 2005).

Here, we focus on the restricted summit endemic treefrog, *Tepuihyla edelcae*, to understand the effects of landscape features on population structure. *Tepuihyla edelcae sensu lato* (Ayarzagüena et al., 1992) has a disjunct distribution atop two formations,

Auyán-tepui and the Chimantá Massif (Figure 3.1). The latter supports many individual tepui summits separated by intermediate elevations and forest (Huber, 1992). Even though populations of *T. edelcae* display few or no morphological differences throughout their distribution (Ayarzagüena et al., 1992; Salerno et al., *in prep*), genetic studies have found the populations atop these two formations to be moderately divergent (2% mtDNA divergence for 12S and ND1 combined) with strong support for reciprocal monophyly of mtDNA lineages (Salerno et al., 2012; Salerno et al., *in review*). Given that the Chimantá massif has an extensive summit area of approximately 800km^2 and enormous landscape complexity (Figure 3.1), the haplotype diversity and population structure in *T. edelcae* is surprisingly low. For four out of five genes sampled, a single common haplotype and a few rare haplotypes are distributed across the entire massif (Salerno et al., *in review*). However, it is unclear whether this pattern reflects a true lack of population structure or a lack of resolution due to the small number of loci.

We generated a Single Nucleotide Polymorphism matrix (SNP) through the massively parallel sequencing technique known as ddRAD (Peterson et al., 2012) from individuals atop four summit localities, three of which inhabit different tepuis on the Chimantá massif. We estimate population structure and divergence between the two main formations (Auyán and Chimantá), as well as among three smaller tepuis on the Chimantá massif. We also contrasted results from Bayesian clustering analyses and multivariate ordination techniques for understanding population structure.

METHODS

Samples, library preparation, and sequencing

We sampled four summit localities within Canaima National Park in the Venezuelan Pantepui, one in Auyán-tepui and three in the Chimantá Massif (Eruoda-tepui, Churi-tepui, and Abakapá-tepui). We sequenced a total of 96 individuals of *Tepuihyla edelcae* from these four summits. Details of sampling localities and museum codes are in Table 3.1. We generated restriction-associated double-digest DNA libraries following the protocol described in Peterson et al. (2012). We extracted DNA with the Viogene Blood and Tissue Genomic DNA Extraction Kit, following the associated protocol. Initial DNA extract normalizations were set for 8ng/μl (for an initial 200ng total to digest). After troubleshooting with several enzyme pairs we chose to perform double digests with the enzymes SphI and MspI and selected a fragment size range of 430–470bp (excluding adaptors). Size selection was done using the Blue Pippin (2% agarose cartridge, Sage Science, Beverly, MA, USA). We used the pre-designed adaptors (P1 flex, and P2) and three PCR primers compatible with the restriction enzymes (Peterson et al. 2012, supplementary materials). The thermocycler protocol for the phusion PCR was the same as in Peterson et al. (2012), with a total of 12 cycles. Sequences were obtained in Illumina Hiseq 2500 as paired-end reads, and our target number of reads was approximately 40 million per pool of 48 individuals.

Bioinformatics and data processing

We processed the raw read data using the *process_radtags* program in the Stacks software pipeline (Catchen et al., 2011, 2013). We used a conservative combination of parameters to discard low-quality reads and clean the data. We also used *denovo_map* in Stacks (Catchen et al., 2011; 2013) for generating the catalog loci for population inference. We repeated *denovo_map* numerous times to examine the effect of input parameters and priors as well as the program's performance with single-end or double-end reads. After analyzing output matrices, we decided to use single-end reads rather than double-end (which greatly reduced percent missing data in final SNP matrix), a minimum stack depth of two, distance of three allowed between stacks (in nucleotide differences), and distance of catalog loci (n) of two. Fst values were calculated using the program *populations*, also part of Stacks (Catchen et al., 2011; 2013), and were lower for n=1 than for n=2 and 3, thus we report the average obtained from all combinations of parameter settings used (Table 3.2). We used the program *populations* to generate the final SNP matrix. After troubleshooting using different parameters, we conservatively chose to keep only the first SNP of each read to avoid read errors as well as linkage issues in downstream analyses. We used only SNPs present in all four populations, and discarded SNPs that had more than 60% missing data per population, which resulted in a matrix with 5,160 SNPs. All final scripts used in STACKS to generate this matrix can be found in the online data repository Dryad (doi:10.5061/dryad.202d4).

In order to examine the effect of missing data on individual assignment among the Chimantá populations, we eliminated all samples with more than 30% missing data in the

matrix and eliminated all SNPs with more than 6% missing data, which resulted in a reduced matrix of 1219 SNPs. Since the goal was to evaluate if population structure and individual assignment improved for the Chimantá populations, this matrix excluded the Auyán population. All matrices used for the analyses can be found in the online data repository Dryad (doi:10.5061/dryad.202d4).

PCA and DAPC analyses

We performed a Principal Components Analysis (PCA) of the unscaled SNP matrix using the *adegenet* package in R (Jombart, 2008). We also performed a Discriminant Analysis of Principal Components (DAPC), which first transforms the data into principal components and then performs a discriminant analysis on the components (Jombart, 2008; Jombart et al., 2010). Since the cumulative variance of principal components did not reach a clear asymptote, we varied the number of PCs retained to examine potential instability issues in the assignment of individuals to a group (Jombart, 2008; Jombart et al., 2010).

We used *find.clusters* (also in *adegenet*) to identify clusters of individuals and compare to the a priori clusters based on known population membership. The command *find.clusters* transforms data into PCs, and then uses a *k*-means algorithm with increasing number of clusters (*k)*. The smallest BIC (Bayesian Information Criterion) value was used to choose the most appropriate *k*. All analyses in adegenet were performed with the larger matrix (5160 SNPs), a subset matrix (1000 SNPs) with all 96 individuals present,

and the matrix where both individuals and SNPs were eliminated to reduce missing data (1219 SNPs, 65 individuals).

STRUCTURE analyses

We estimated genetic clusters using STRUCTURE v2.3 (Pritchard et al., 2000). To reduce computation time, we used the optimal number of clusters obtained in *adegenet* as a guide to perform analyses from k=1 to k=8, with four iterations each. We performed analyses of the full dataset (four populations) assuming no admixture and uncorrelated allele frequencies, and assuming admixture and correlated allele frequencies (Falush et al., 2003). We then repeated the analyses using pre-defined populations as priors. We performed an analysis of the three Chimantá populations, excluding Auyán, under the admixture model using correlated allele frequencies. We determined the optimal number of clusters using the Evanno method in Structure Harvester (Evanno et al., 2005; Earl and VonHoldt, 2012). We also used CLUMPP (Jakobsson and Rosenberg, 2007) to permute STRUCTURE runs into a single output. We generated graphs using DISTRUCT (Rosenberg, 2004). The analyses in STRUCTURE were done using both the subset matrix with all individuals present (1000 SNPs, 96 individuals) and the matrix with culled missing data (1219 SNPs, 65 individuals).

RESULTS

PCA and DAPC analyses

The PCA analysis revealed an extreme outlier in the Auyán population (Figure S3.1); this was eliminated from the matrix due to possibility of contamination. When using the full SNP matrix, the PCA revealed the presence of four genetic clusters corresponding to the four original populations (Figure 3.2). The first PC, which made up most of the cumulative variance, separated the Auyán population from all three Chimantá populations. The second PC separated the three Chimantá populations into three clusters corresponding to the *a priori* populations. However, some individuals either fell within the 95% ellipses of other populations or outside the ellipses near the origin point (0, 0) of the biplots (Figures 3.2b and 3.3b).

The DAPC assignment proportions and clustering patterns vary depending on the number of PCs retained, most notably for the clustering and separation of the Abakapá, Churi, and Eruoda populations of Chimantá (Figure 3.3). When the number of PCs retained was varied from 10 to 60, assignment proportions for Churi increased from 85% to 100%, and the assignment proportions of Eruoda decreased from 100% to 97%. The Auyán and Abakapá assignment proportions remained the same (100% and 91% respectively). Given that retaining too many PCs may result in instability of assignment proportions (Jombart et al. 2010), we kept only 30 PCs, which corresponded to the peak of overall assignment proportion (95.6%). For this number of PCs the assignment proportions were 100% for Auyán and Abakapá, 85% for Churi, and 97.4% for Eruoda.

The clustering function (*k*-means algorithm in *adegenet*) identified k=5 as the optimal number of clusters given 30 retained PCs. Figure 3.5 shows that at k=4, five

individuals from Abakapá and four from Churi are incorrectly assigned to the Eruoda cluster, while those from Eruoda and Auyán are always correctly assigned. For *k*=5 the Eruoda cluster is separated into two sub-clusters: one unique to Eruoda, and one with individuals shared with both Abakapá and Churi. When using the smaller matrix but without reducing missing data or individuals (1000 SNPs), the results were nearly identical to the full matrix for all *adegenet* analyses.

The set of analyses that excluded the Auyán population but retained all individuals of the other three populations showed mixed results. The PCA and cluster analyses showed only a slight effect of the removal of Auyán (Figures 3.2b and 3.3b). The Principal Components and the Linear Discriminants (DAPC) are, not surprisingly, much more informative of the distances and discrimination among the three Chimantá clusters if Auyán is excluded, due to the first two PCs being almost equally important to explain variation (Figure 3.2b). However, the correct assignment proportions of the clusters in the DAPC analysis decreased substantially for the Abakapá population, from 100% to 76.2%. The assignment probabilities remained identical regardless of number of retained PCs (76.2% for Abakapá, 85% for Churi, and 100% for Eruoda). The clustering algorithm unambiguously supported the same number of original populations as clusters (*k*=3) and retained the incorrect assignment of several Abakapá and Churi individuals to the Eruoda population (Figure 3.5b).

When using the matrix that eliminated missing data both by removing taxa and SNPs, no individuals fell within or near the biplot origin, and thus there was complete

differentiation among populations in both the PCA and the DAPC (Figure 3.4). The DAPC assignment proportions were 100% for all three populations.

STRUCTURE analyses

Analyses using the admixture and no-admixture models with the subset matrix (all samples) yielded nearly identical results, so we only present the admixture analyses (Figure 3.6a). The Evanno method (Evanno et al., 2005; Earl and vonHoldt, 2012) favored k=2 for all analyses using the full dataset (four populations). The analyses that did not use population assignment as priors assigned two individuals of the Chimantá massif to Auyán (Figure S3.2). However, analyses that used population membership as priors inferred only some level of admixture in the two Churi individuals that were also assigned to the Auyán cluster when not using the population priors (Figure 3.6a).

The analyses excluding the divergent population (Auyán) favored k=3, which corresponds to the original number of localities (Figure 3.6b). All three populations showed moderate to high levels of admixture for a few individuals. However, when analyzing the matrix with excluded taxa (to minimize missing data to less than 30% per taxa), we observe that all individuals that had moderate levels of admixture inferred were eliminated in the second matrix. Thus, in this analysis, which also favored k=3, all individuals were unequivocally assigned to their population of origin; negligible amounts of admixture were inferred for only a few individuals (Figure 3.6c).

The Fsts obtained from the Stacks *populations* analyses (Catchen et al., 2011; 2013) varied depending on prior settings. We report average and standard deviations

obtained from all ten combinations of parameter settings (Table 3.2). The highest Fsts were for the Auyán-Chimantá population pair comparisons, and the lowest Fst were between Eruoda and Churi, which are the geographically closest populations.

DISCUSSION

Cluster inference, individual assignment, and comparison between multivariate and Bayesian clustering

Inferring genetic clusters of populations is not always straightforward. Many clustering methods rely on Bayesian algorithms, which require many prior assumptions, have complex evolutionary models and long computation times (Pritchard et al., 2000; Jombart et al., 2010; Kalinowski, 2011). Multivariate ordination methods such as PCA and DAPC do not rely on complex evolutionary models, which greatly reduces the computational time and the potential biases introduced by priors. However, the ordination methods differentially weight variables (SNPs), and in doing so may reduce the effect of less variable SNPs and rare alleles (Jombart, 2008; Jombart et al., 2010). Multivariate analyses are considered a good complement to Bayesian clustering methods because the methods are inherently different and because ordination methods allow visualization of genetic distances in space (Jombart et al., 2010; Kalinowski, 2011).

The DAPC correctly assigned 95.6% of individuals to their original clusters; the only incorrectly assigned individuals were those that fall near the biplot origin point (Figure 3.3). These results are concordant both with geographical distributions of sampling localities and with previous data (Salerno et al., *in review*). Previous

phylogenetic studies found the two main clusters, Auyán and Chimantá, to be reciprocally monophyletic with 100% posterior probability, regardless of methods or datasets. The two clusters had approximately 2% pairwise mtDNA sequence divergence (Salerno et al. 2012; Kok et al., 2012; Jungfer et al., 2013; Salerno et al., *in review*). Also, each cluster is found on different geological formations with intervening lowlands having different flora and climate, and in which no close relatives of these lineages have been found. Given this, the PCA and DAPC analyses are consistent with expectations of a high level of differentiation between the Auyán and Chimantá groups.

Analysis of the matrix that excluded the taxa with high levels of missing data (>30%) showed that all individuals were correctly assigned to their original populations, and no individuals fell within the origin of the biplot, suggesting that missing data, and not genetic exchange, was causing this pattern of shared haplotypes among the three populations.

The *k*-means clustering analysis (*adegenet*) favored five clusters (based on BIC) if all four populations were analyzed, and three clusters if the most divergent population, Auyán, was excluded. In the optimal clustering scheme (k=5), four of the five clusters correspond to the correctly classified individuals. The fifth cluster includes a few misclassified individuals from Eruoda, Abakapá, and Churi (Figure 3.5A). Comparison of these results with the PCA and DAPC plots (Figures 3.2 and 3.3) confirm that the misclassified individuals correspond to those located near the biplot origin, suggesting that these individuals are genetically similar to each other. However, these taxa were eliminated in the smaller matrix (with reduced missing data), and the clusters were

correctly inferred after their exclusion. Thus, we interpret this to mean that the few misclassified individuals are due to uncertainty from high levels of missing data. That all individuals are correctly classified in the matrix with reduced missing data reflects the genetic distinctiveness of each population.

The STRUCTURE analyses that included all taxa inferred only two clusters, regardless of whether the admixture model was used and whether population assignments were used as priors. The two clusters comprised the Auyán population and the three combined Chimantá populations (Figure 3.6a). The STRUCTURE analyses of the three Chimantá populations alone (Figure 3.6b) agreed with the multivariate analyses in supporting three clusters, consistent with the original number of sampled populations. Thus, both the multivariate and Bayesian approaches yielded largely correct classifications of individuals when Auyán was excluded.

In STRUCTURE analyses in which localities were used as priors, and regardless of the models used, the individual assignments generally agreed with other analyses, biogeography, and previous data (Salerno et al., *in revsion*), since those two incorrectly assigned individuals were inferred simply to have some level of ancient admixture. However, in both analyses in which population priors were not used, two individuals of Churi were assigned to Auyán with high probability, regardless of *k* (Figure S3.2). This result contradicts the PCA and DAPC results, as well as previous studies (Salerno et al., 2012), which found these two lineages to be divergent, well-supported monophyletic groups.

This conflicting result may reflect problems associated with STRUCTURE (Aurelle et al., 2011; Kalinowski et al., 2011). In a simulation study Kalinowski (2011) found that STRUCTURE incorrectly inferred clusters when there were substantial differences in divergence among groups (clusters) or sample sizes. Given that STRUCTURE sorts individuals into Hardy-Weinberg populations (Pritchard et al., 2000), and by doing so maximizes the overall likelihood of the inference instead of the individual likelihoods of each cluster, it may create one homogeneous cluster and another highly variable "wastebasket" cluster when k=2 (Kalinowski, 2011). In fact, comparison of our STRUCTURE analyses to pairwise population Fst values (Table 3.2) showed that the Fst values do not support these incongruent results, given that the lowest Fst value for the divergent populations is between Auyán and Eruoda (0.342 ± 0.0441 S.D.), rather than between Auyán and Churi (0.399 ± 0.0444) that have admixed individuals (Table 3.2). Thus, we are confident that the analyses using populations as priors are more representative of the true clustering pattern between Auyán-tepui and the Chimantá Massif.

On the utility of SNPs obtained through massively parallel sequencing techniques for analyses of demographic history

The results Salerno et al. (*in revision*), where only a few loci were used, together with this current study, suggest a recent demographic expansion of *T. edelcae* atop the Chimantá Massif. However, Salerno et al. (*in revision*) used only four unlinked loci for its estimates of demographic expansion. It is well established that accurate estimation of

demographic histories requires a large number of unlinked loci to observe genome-wide effects (Brumfield et al. 2003; Luikart et al. 2003; Seeb et al. 2011; Excoffier et al. 2013). Individual genes have unique histories and different mutation rates; moreover, estimates of demographic histories may be biased by non-neutral loci (Brumfield et al. 2003; Luikart et al. 2003; Gattepaille et al. 2013). Neutral loci that are unlinked to areas under selection should be similarly affected by demography throughout the genome (Brumfield et al. 2003; Luikart et al. 2003). Thus, sampling strategies that take a whole-genome approach, such as microsatellites, Single Nucleotide Polymorphisms (SNPs), and Associated Fragment Length Polymorphisms (AFLPs), are ideal for demographic estimates (Brumfield et al. 2003). Although it is debated, most authors agree that SNPs are superior markers for demographic inference (Brumfield et al. 2003).

Most traditional approaches for estimating demographic history have been developed for model organisms, due to the much higher availability of data for these species (Gattepaille et al. 2013). Thus, many analyses rely on a well-mapped and annotated genome for inferring linkage (Gattepaille et al., 2013). These methods, sometimes referred to as the class two summary statistics, are very powerful statistical tools for detecting demographic expansion, but are also highly sensitive to recombination (Gattepaille et al., 2013).

Traditionally, most of the analyses and software available for estimating demographic histories with SNPs had to account for non-trivial issues regarding ascertainment bias (Garvin et al. 2010; Seeb et al. 2011; Korneliussen et al. 2013), which refers to bias inherent in SNP chip detection approaches. In these approaches, most

researchers would screen a few individuals and utilize those previously identified SNPs for the rest of the individual samples. This approach is problematic in that rare alleles and unique haplotypes, which are more recently derived in the gene pool, are more likely to be missed in this initial screening phase, and thus will bias the coalescent reconstruction of the demographic history towards overestimation of migration rates (Brumfield et al. 2003). Many of the approaches that are in use today account for this type of ascertainment bias (Garvin et al. 2010; Seeb et al. 2011; Korneliussen et al. 2013).

With the advent of more affordable high-throughput sequencing techniques, approaches that were traditionally reserved only for model organisms with available genomes are now possible with non-model organisms having little to no available data (Garvin et al. 2010; Seeb et al. 2011). Most of these are based on genome reduction and alignment of redundant reads, such as is possible with restriction site-associated DNA markers (Miller et al. 2007; Baird et al. 2008; Seeb et al. 2011; Hohenlohe et al. 2011; Peterson et al. 2012). With these massively parallel sequencing techniques, which allow simultaneous discovery and genotyping of thousands of SNPs, previous issues with ascertainment bias for SNPs are eliminated, although other potential biases are introduced during data gathering (Seeb et al. 2011; Korneliussen et al. 2013).

In the absence of reference genomes, the most common analyses of genomic data obtained through massively parallel sequencing techniques are based on the frequency spectrum, or the distribution of frequencies of segregating sites on a sample of alleles (Braverman et al. 1995). These are also referred to as class 1 summary statistics, of which the most commonly used is the well-known Tajima's D statistic (Tajima 1989;

Braverman et al. 1995). Site Frequency Spectrum (SFS) statistics are extremely sensitive to loci or regions of the genome that are prone to selective effects, and thus these regions should be excluded from analyses (Luikart et al. 2003). Given that for non-model organisms with no reference genome we have no information whether the SNPs are in coding regions, nor information on linkage among SNPs, it is crucial to perform some form of screening for areas under selection, including both regions under direct selection and areas prone to hitchhiking (Luikart et al. 2003). The most common approach is to scan for outlier loci, where outliers in both high and low Fst values are excluded from the dataset to eliminate effects of selection (Luikart et al. 2003).

It also may be desirable to make inferences about population size as a function of time. The most commonly used methods are Bayesian Skyline plots and Extended Bayesian Skyline plots, which yield non-parametric estimates of population size changes in mitochondrial and multiple unlinked loci, respectively (Drummond et al. 2002; Heled and Drummond 2008). However, these methods are not yet well established for SNP data. Thus, some methods have been specifically designed, based on more complex statistical analyses of the SFS and AFS (Allele Frequency Spectrum), for genome-wide and low-coverage SNP data. These methods can account for several types of ascertainment bias and estimate the population size through composite likelihood (Gutenkunst et al. 2009; Naduvilezhath et al. 2011; Excoffier et al. 2013; Korneliussen et al. 2013), Bayesian, (Korneliussen et al. 2013), and Approximate Bayesian Computation methods (Theunert et al. 2012). However, much needs to be done in terms of establishing the potential biases introduced by these newer techniques and thus adequately modify the approaches and

available software for estimating population size using data obtained with massively parallel sequencing techniques.

Tepuis as sky islands: significance of population structure within the Chimantá Massif

Salerno et al. (*in review*) found unexpectedly low levels of genetic diversity and population structure within and among the three populations of *T. edelcae* atop the Chimantá Massif. Given the level of landscape complexity, steepness of the tepui walls (Figure 3.2), and habitat heterogeneity (Huber, 1992) atop the massif, we expected that a distance of more than 30km, which is observed among two of the three population pairs, would be sufficient to generate detectable population structure, even with low-throughput (Sanger) sequencing. However, genetic diversity was much lower among the combined Chimantá populations than within the single locality atop Auyán, suggesting that a very recent bottleneck occurred in the Chimantá Massif populations (Salerno et al., *in review*).

Kok et al. (2012) claimed that genetic diversity in ND2 the gene atop the tepui summits was low. Although they did not explicitly estimate genetic diversity, they found surprisingly low levels of genetic divergence between neighboring tepui summits. Their result is similar to that of Salerno et al. (*in review*) for the Chimantá Massif, who used a larger number of genes (five). Together, the two results suggest that a recent event drastically reduced the effective population sizes of several high-elevation tepui summits such as the Eastern tepui chain (Kok et al., 2012).

Not surprisingly given the differences in geographic distance, mtDNA divergences within the Chimantá massif are very low compared to divergences between

Chimantá and Auyán (Figure 3.2A). On the other hand, the genetic structure within the Chimantá Massif as assessed by SNPs is still sufficient that all individuals are correctly assigned to their population (when excluding taxa with high levels of missing data). This suggests very rapid accumulation of variation unique to each population, likely driven by a combination of reduced gene flow and the effects of drift. The high levels of genetic structure in the SNP dataset, evident in both the multivariate and Bayesian analyses, suggest landscape complexity is driving population isolation among these summits.

The summits of the Chimantá Massif and Auyán-tepui are highly pristine ecosystems due to their high inaccessibility and isolation, and as such we can be certain that human impacts and modified landscapes are not responsible for generating any of the population structure. Further studies should incorporate direct measures of population connectivity, such as least cost path analyses (Storfer et al., 2007; Spear et al., 2010), to explicitly connect landscape complexity and genetic divergence atop these massifs.

Table 3.1: Specimens and localities. TNHC-FS = Texas Natural History Collection Field Series, University of Texas; PS = Patricia Salerno (specimens deposited in MHNLS, Museo de Historia Natural La Salle).

General locality	N	Field numbers	Coordinates	Specific locality
Auyán	11	TNHC-FS05824–05836	5°46.599'N 62°32.251'W	Campamento el Oso, Auyán-tepui, Canaima National Park, Bolívar State, Venezuela
Abakapá	21	PS365–371, PS387–404	5°11.497'N 62°18.939'W	Abakapá-tepui, Chimantá Massif, Canaima National Park, Bolívar State, Venezuela
Churi	20	PS330–339, PS345–346, PS357–364	5°15.257'N 62°00.472'W	Churi-tepui, Chimantá Massif, Canaima National Park, Bolívar State, Venezuela
Eruoda	38	PS410–411, PS437, PS446–466, PS_R_01–17	5°22.525'N 62°05.674'W	Eruoda-tepui, Chimantá Massif, Canaima National Park, Bolívar State, Venezuela

Table 3.2: Fst values calculated using *populations* (Stacks). Average and standard deviations are shown for all Fst values using ten different combinations of de_novo priors. Numbers in bold represent comparisons among the Auyan-tepui and Chimantá Massif formations, while the others represent comparisons within Chimantá.

	Abakapá	Churi	Eruoda
Auyán	**0.404 ± 0.0475**	**0.399 ± 0.0444**	**0.342 ± 0.0441**
Abakapá	-	0.131 ± 0.0047	0.079 ± 0.0023
Churi	-	-	0.077 ± 0.0025

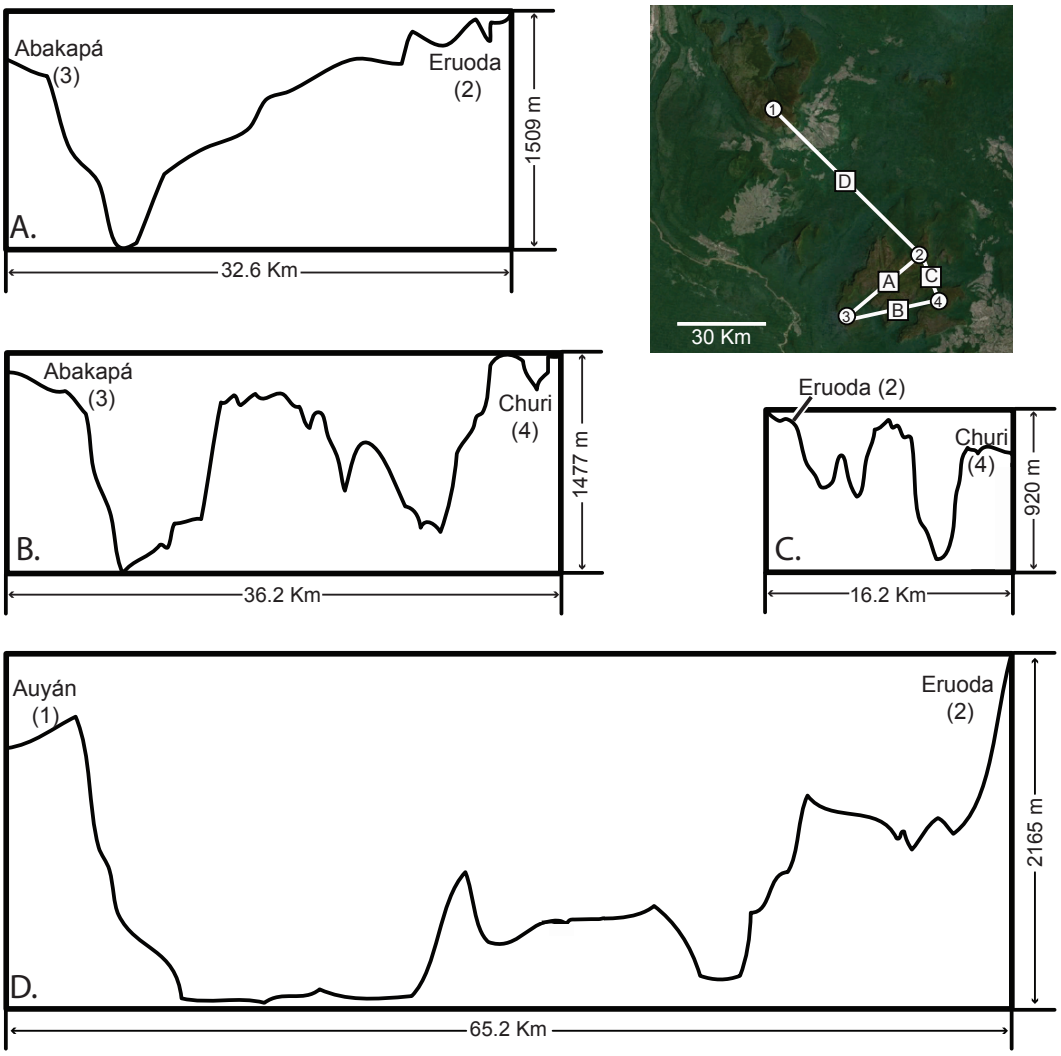

Figure 3.1. Map of sampling localities and elevation profiles of transects (A–D), where (1) is the single Auyan-tepui population, and (2), (3), and (4) are the three populations within the Chimantá Massif: Eruoda, Churi, and Abakapá, respectively.

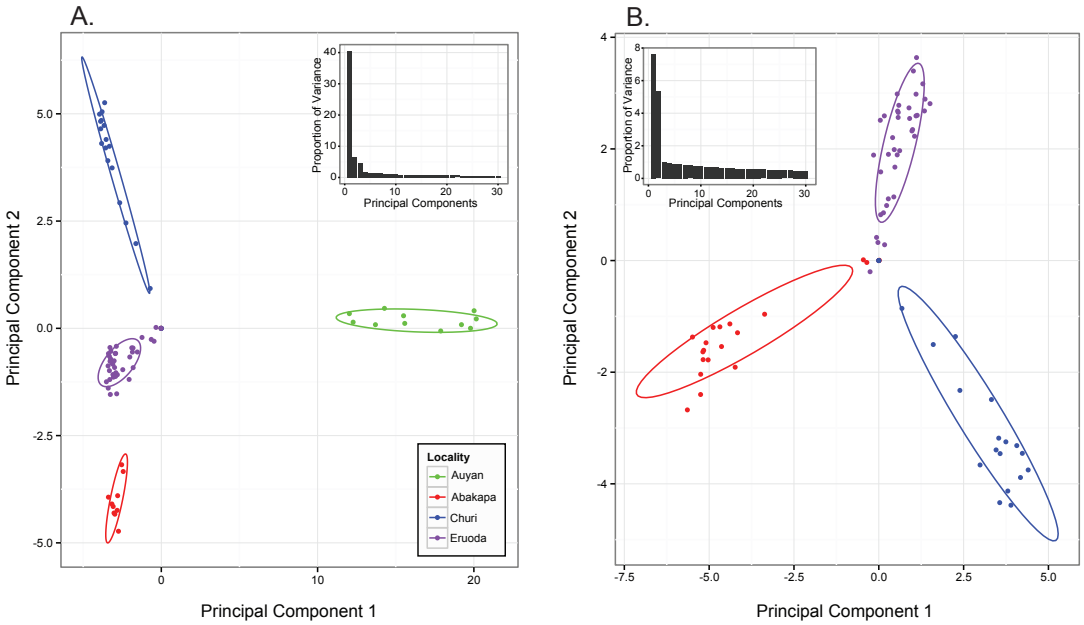

Figure 3.2. Principal Components analyses including all four populations (A), and excluding the divergent population, Auyán (B). PC eigenvalues are shown as inserts; bars represent percent contribution to the variance of each principal component. Only PC1 and 2 are plotted.

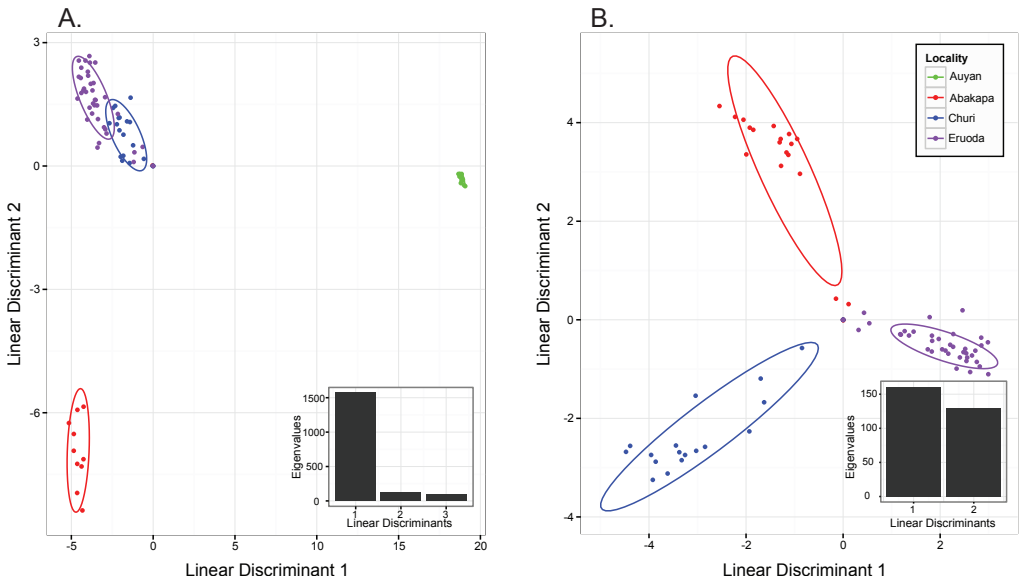

Figure 3.3. Discriminant Analysis of Principal Components for all populations (A), and excluding the divergent population, Auyán (B). DA eigenvalues are shown as inserts; bars represent contribution of each linear discriminant to the among-group variation.

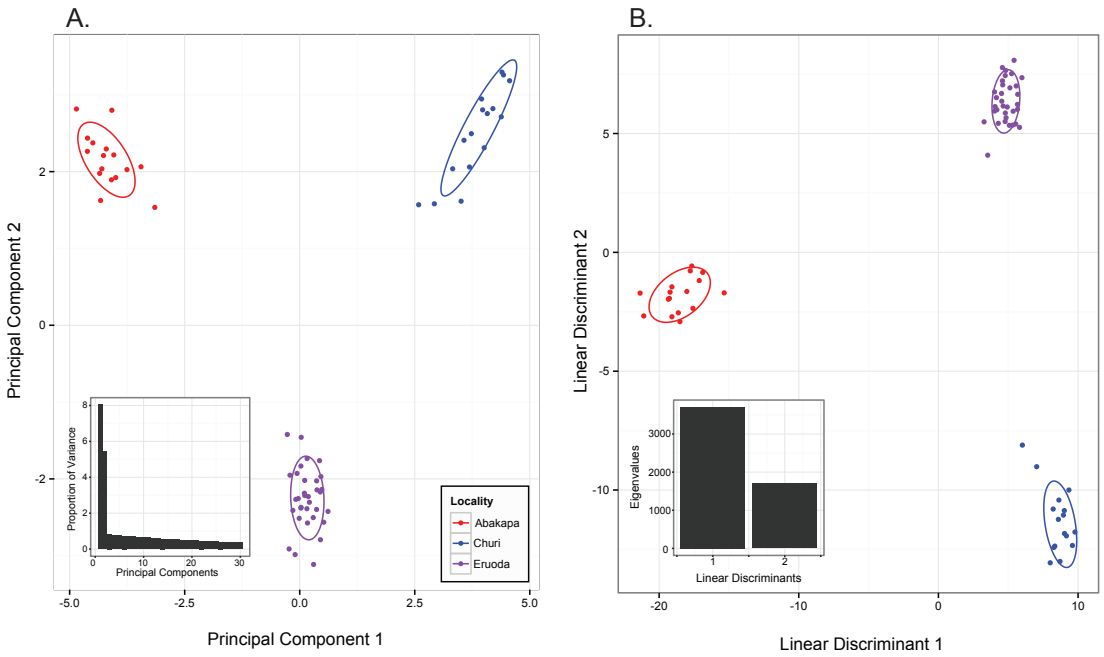

Figure 3.4: Principal components analysis (A) and discriminant analysis of principal components (B) for the matrix that excludes taxa with >30% missing data and SNPs with >6% missing data.

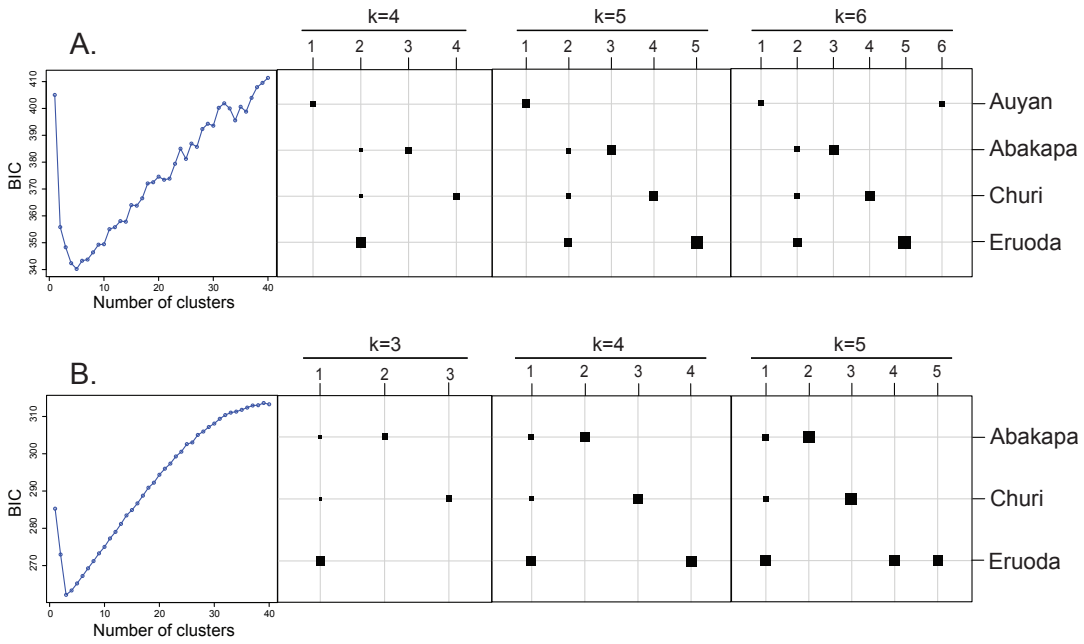

Figure 3.5. *k*-means clusters for all four populations (*adegenet* analysis) (A) and excluding the divergent population, Auyán (B). Inferred clusters are on top (labeled with numbers) and original clusters are on the right (labeled with population names). Bayesian Information Criterion (BIC) values are plotted for k=1 through k=40. The favored number of clusters was five for (A) and three for (B).

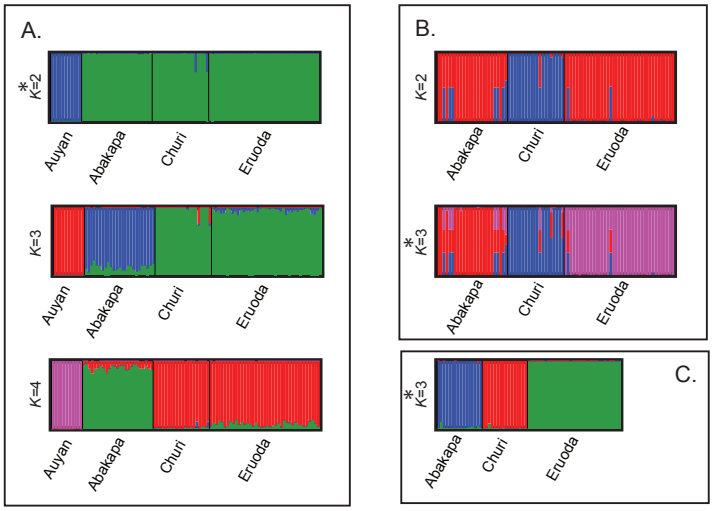

Figure 3.6. Results from STRUCTURE analysis using admixture and population assignment as priors. (A) All four populations, (B) Three populations, excluding Auyán, the divergent population, and (C) Three populations, excluding Auyán, using the matrix that excluded samples with >30% missing data. Each vertical bar represents an individual; the height of the bar represents the probability of its assignment to the population. Asterisks represent favored *k* for each analysis.

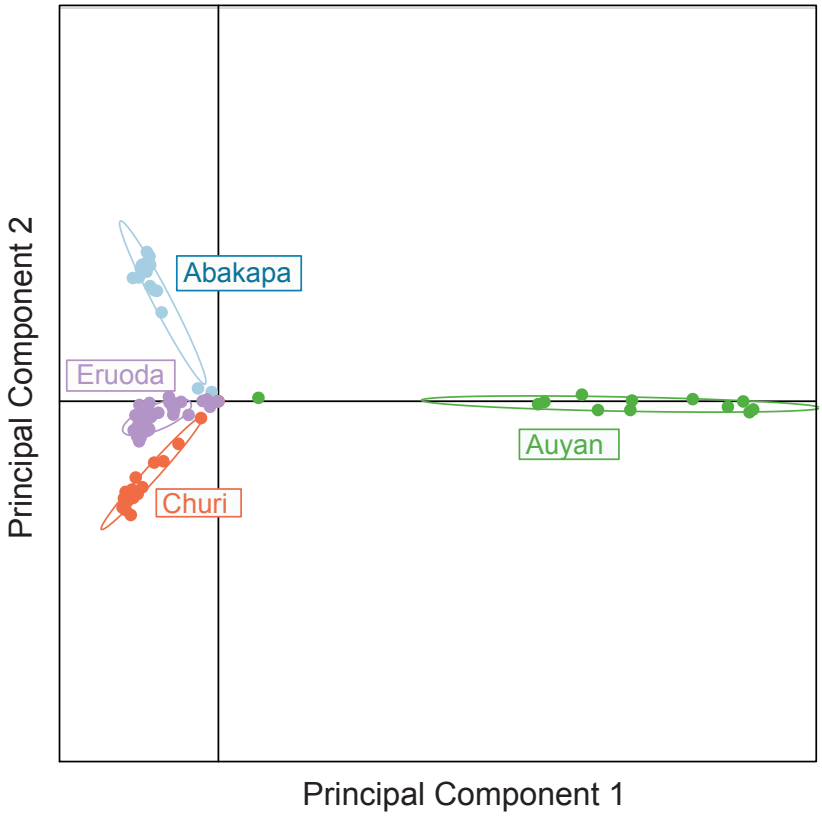

Figure S3.1: Principal Components analysis of all four populations before exclusion of outlier in the Auyán population.

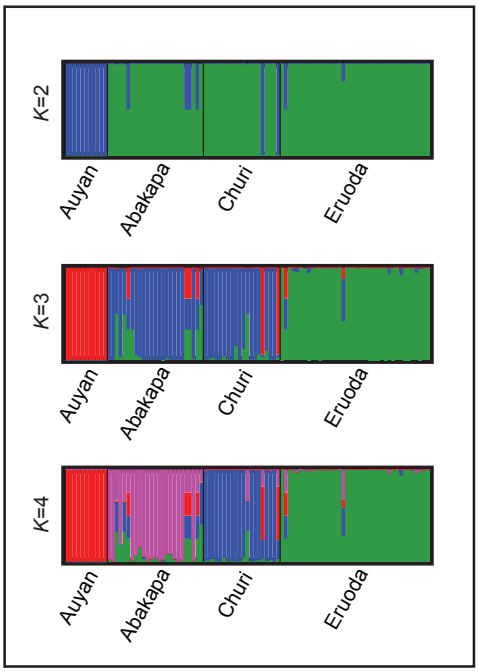

Figure S3.2: Results from STRUCTURE analysis for all four populations using admixture but without population assignment as priors. Analysis favored $k = 2$.

References

Alonso, R., Crawford, A.J., and Bermingham, E. 2012. Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes. Journal of Biogeography 39: 434-451.

Austin, J.D., Dávila, J.A., Lougheed, S.C., Boag, P.T. 2003. Genetic evidence for female-biased dispersal in the bullfrog, *Rana catesbeiana* (Ranidae). Molecular Ecology 12: 3165–3172.

Aurelle, D., Ledoux, J.-B., Rocher, C., Borsa, P., Chenuil, A., and Féral, J.P. 2011. Phylogeography of the red coral (*Corallium rubrum*): inferences on the evolutionary history of a temperate gorgonian. Genetica 139: 855–869.

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge.

Avise, J.C. 2009. Phylogeography: retrospect and prospect. J. Biogeogr. 36: 3–15.

Ayarzagüena, J., Señaris, J.C., and Gorzula, S. 1992. The group *Osteocephalus rodriguezi* from the highlands of the Venezuelan Guayana: description of five new species. Memoria de la Sociedad de Ciencias Naturales La Salle 52: 113–142.

Baird NA, Etter PD, Atwood TS et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3, e3376.

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13: 729–744.

Ballard, J.W.O., and Kreitman, M. 1995. Is mitochondrial DNA a strictly neutral marker? Trends in Ecology and Evolution 10: 485–488.

Barrio-Amorós, C. L. 2004. Amphibians of Venezuela. Systematic list, distribution and references, an update. Rev. Ecol. Lat. Amer. 9: 1–48.

Bazin, E., Glémin, S., and Galtier, N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312: 570–572.

Bech, N., Boissier, J., Drovetski, S., and Novoa, C. 2009. Population genetic structure of rock ptarmigan in the 'sky islands' of French Pyrenees: implications for conservation. Animal Conservation 12: 138–146.

Bell, C.J., Gauthier, J.A., and Bever, G.S. 2010. Covert biases, circularity, and apomorphies: A critical look at the North American Quaternary Herpetofaunal Stability Hypothesis. Quat. Int. 217: 30–36.

Berry, P.E. and Riina, R. 2005. Insights on the diversity of the Pantepui flora and the biogeographic complexity of the Guayana Shield. Biologiske Skrifter 55:145–167.

Bonaccorso, E., and Guayasamin, J.M. 2013. On the origin of Pantepui montane biotas: A perspective based on the phylogeny of *Aulacorhynchus* toucanets. PLoS ONE 8, 0067321. doi:10.1371/journal.pone.0067321.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T.G., Wu, C-H., Xie, D., Suchard, M.A., Rambaut, A., and Drummond, A.J. 2014. BEAST2: A software platform for Bayesian evolutionary analysis. PLoS Computational Biology 10, e1003537.

Braverman, J.M., Hudson, R.R., Kaplan, N.L., Langley, C.H., and Stephan, W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics 140: 783–796.

Briceño, H. O., C. Schubert, and J. Paolini. 1990. Table-mountain geology and surficial geochemistry: Chimantá Massif, Venezuelan Guayana shield. J. South Amer. Earth Sci. 3: 179–194.

Briceño, H. O. and Schubert, C. 1990. Geomorphology of the Gran Sabana, Guayana Shield, southeastern Venezuela. Geomorphology 3: 125–141.

Brown, J.H., and Lomolino, M.V. 1998. Biogeography. 2nd ed. Sinauer Associates, Sunderland, MA.

Brumfield, R.T., Beerli, P., Nickerson, D.A., and Edwards, S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution 18: 249–256.

Carnaval, A. C. and Moritz, C. 2008. Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. Journal of Biogeography 35: 1187–1201.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics 1: 171–182.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., and Cresko, W. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22: 3124–3140.

Chapman, F.M. 1931. The upper zonal bird-life of Mts. Roraima and Duida. Bulletin of the American Museum of Natural History 63: 1–135.

Crews, S.C. and Gillespie R.G. 2010. Molecular systematics of *Selenops* spiders (Araneae: Selenopidae) from North and Central America: implications for

Caribbean biogeography. Biological Journal of the Linnaean Society 101: 288–322.

Dávalos, L. M. 2004. Phylogeny and biogeography of Caribbean mammals. Biological Journal of the Linnaean Society 81: 373–394.

DeChaine, E.G., and Martin, A.P. 2005. Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains. American Journal of Botany 92: 477–486.

Delport, W., Poon, A., Frost, S.D.W., and Kosakovsky Pond, S.L. 2010. Datamonkey: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26 2455-2457.doi:10.1093/bioinformatics/btq429.

Dowling, D.K., Friberg, U., and Lindell, J. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. Cell 23: 546–554.

Drummond, A.J., and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7: 214.

Duellman, W.E. 1999. Patterns of distribution of amphibians: a global perspective. Johns Hopkins University Press, Baltimore, MD.

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., and Solomon, W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161: 1307–1320.

Earl, D.A., and VonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Evans, B.J., Kelley, D.B., Tinsley, R.C., Melnick, D.J., and Cannatella, D.C. 2004. A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. Molecular Phylogenetics and Evolution 33: 197–213.

Excoffier, L., and Lischer, H.E.L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and Foll, M. 2013. Robust demographic inference form genomic and SNP data. PLoS Genetics 9: e1003905.

Falush, D., Stephens, M., and Pritchard, J.K. 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics 164: 1567–1587.

Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A., Wheeler, W.C. 2005. Systematic review of the frog family Hylidae, with special reference to Hylinae: phylogenetic analysis and taxonomic revision. Bulletin of the American Museum of Natural History 294: 1–240.

Fouquet, A., Gilles, A., Vences, M., Marty, C., Blanc, M., and Gemmel, N. 2007. Underestimation of species richness in Neotropical frogs revealed by mtDNA analyses. PLoS ONE 2 e0001109. doi:10.1371/journal.pone.0001109.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915–925.

Galbreath, K.E., Hafner, D.J., and Zamudio, K.R. 2009. When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American Pika, *Ochotona princeps*). Evolution 63: 2848–2863.

Galtier, N., Nabholz, B., Glémin, S., and Hurst, G.D.D. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18: 4541–4550.

Garvin, M.R., Saitoh, K., and Gharrett, A.J. 2010. Application of single nucleotide polymorphisms to non-model species: a technical review. Molecular Ecology Resources 10: 915–934.

Gattepaille, L.M., Jakobsson, M., and Blum, M.G.B. 2013. Inferring population size changes with sequence and SNP data: lessons from human bottlenecks. Heredity 110: 409–419.

Gibbs, A.K., and Barron, C.N. 1993. The Geology of the Guiana Shield. Oxford University Press, Oxford, England.

Givnish, T.J., Sytsma, K., Smith, J.F., Hahn, W.J., Benzing, D.H., and Burkhardt, E.M. 1997. Molecular evolution and adaptive radiation in *Brocchinia* (Bromeliaceae: Pitcairnioideae) atop tepuis of the Guayana Shield. In: Givnish, T.J., Sytsma, K.J. (Eds.), Molecular Evolution and Adaptive Radiation. Cambridge University Press, Cambridge, pp. 259–312.

Givnish T.J., Evans, T.M., Zjhra, M.L., Patterson, T.B., Berry, P.E., and Sytsma, K.J. 2000. Molecular evolution, adaptive radiation, and geographic diversification in the

amphiatlantic family Rapateaceae: evidence from ndhF sequences and morphology. Evolution 54: 1915–1937.

Givnish, T.J., Millam, K.C., Evans, T.M., Hall, L.C., Pires, J.C., Berry, P.E., and Sytsma, K.J. 2004. Ancient vicariance or recent long-distance dispersal? Inferences about phylogeny and South American-African disjunctions in *Rapateaceae* and *Bromeliaceae* based on ndhF sequence data. International Journal of Plant Sciences 165: S35–S54.

Givnish, T.J., Barfuss, M.H.J., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P.A., Jabaily, R.S., Crayn, D.M., Smith, J.A., Winter, K., Brown, G.K., Evans, T.M., Holst, B.K., Luther, H., Till, W., Zizka, G., Berry, P.E., and Sytsma, K.J. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: insights from an eight-locus plastid phylogeny. American Journal of Botany 98: 872– 895.

Gorzula, S., and Señaris, J.C. 1998. Contribution to the herpetofauna of the Venezuelan Guayana I. A database. Scientia Guaianae 8: 1–269.

Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., and Bustamante, C.D. 2009. Inferring the joint demographic history of multiple populations form multidimensional SNP frequency data. PLoS Genetics 5: e1000695.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165: 131–137.

Hedges, S.B. 2006. Paleogeography of the Antilles and origin of West Indian terrestrial vertebrates. Ann. Miss. Bot. Gard. 93: 231–244.

Heinicke, M.P., Duellman, W.E., Trueb, L., Means, D.B., MacCulloch, R.D., and Hedges, S.B. 2009. A new frog family (Anura: Terrarana) from South America and an expanded direct-developing clade revealed by molecular phylogeny. Zootaxa 2211: 1–35.

Heled, J., Drummond, A.J. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27: 570–580.

Hohenlohe, P., Amish, S., Catchen, J., Allendorf, F., and Luikart, G. 2011. Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Molecular Ecology Resources 11: 117–122.

Holman, J.A. 1998. Pleistocene amphibians and reptiles in Britain and Europe. Oxford Univ. Press, New York.

Holman, J.A. 2003. Fossil frogs and toads of North America. Indiana Univ. Press, Bloomington, IN.

Hoogmoed, M.S. 1979. The herpetofauna of the Guianan region. Pp. 241–280 *in* W.E. Duellman, ed. The South American herpetofauna: its origin, evolution and dispersal Monogr. Mus. Nat. Hist. Univ. Kansas 7: 1–485.

Hooghiemstra, H., and Van der Hammen, T. 2004. Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. Philosophical Transactions of the Royal Society of London B 359: 173–181.

Hoorn, C. 1993. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. Palaeogeography Palaeoclimatology Palaeoecology 105: 267–309.

Huber, O. 1987. Consideraciones sobre el concepto de Pantepui. Pantepui 2:2–10.

Huber, O. 1988. Guayana highlands versus Guayana lowlands, a reappraisal. Taxon 37: 595–614.

Huber, O. 1992. El Macizo de Chimantá, Escudo de Guayana, Venezuela: Un ensayo ecológico tepuyano. Oscar Todtmann Editores, Caracas, 343 pp.

Huber, O. 1995. Geographical and physical features. *In*: Steyermark, J.A., Berry, P.E. and Holst, B.K. (Eds.), Flora of the Venezuelan Guayana. Vol. 1. Introduction, pp. 1–61. Missouri Botanical Garden and Timber Press, Portland.

Huber, O. 2006. Herbaceous ecosystems on the Guayana Shield, a regional overview. J. Biogeogr. 33: 464–475.

Huber, O., Febres, G., and Arnal, H. 2001. Ecological guide to the Gran Sabana. Ecograph Proyectos y Ediciones, Caracas, 192 pp.

Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. Proceedings of the National Academy of Sciences USA 103: 10334–10339.

Iturralde-Vinent, M.A., and MacPhee, R. 1999. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. Bulletin of the American Museum of Natural History 238: 1–95.

Jaffé, K., Lattke, J., and Perez Hernández, R. 1993. Ants on the tepuis of the Guiana Shield: a zoogeographic study. Ecotropicos 6: 22–29.

Jakobsson, M., and Rosenberg, N. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Jetz, W., Rahbek, C., and Colwell, K. 2004. The coincidende of rarity and richess and the potential signature of history in centres of endemism. Ecology Letters 7: 1180–1191.

Johnson, N.K., and Cicero, C. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. Evolution 58: 1122–1130.

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403–1405.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11: 94.

Jungfer, K.H., Faivovich, J., Padial, J., Castroviejo-Fisher, S., Lyra, M.M., Berneck, B., Iglesias, P.P., Kok, P.J.R., MacCulloch, R.D., Rodrigues, M.T., Verdade, V.K., Torres Gastello, C.P., Chaparro, J.C., Valdujo, P.H., Reichle, S., Moravec, J., Gvozdik, V., Gagliardi-Urrutia, G., Ernst, R., De la Riva, I., Means, D.B., Lima, A.P., Señaris, J.C., Wheeler, W.C., and Haddad, C.F.B. 2013. Systematics of spiny-backed treefrogs (Hylidae: *Osteocephalus*): an Amazonian puzzle. Zoologica Scripta 42: 351–380.

Kalinowski, S.T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106: 625–632.

Karl, S.A., Toonen, R.J., Grant, W.S., Bowen, B.W., 2012. Common misconceptions in molecular ecology: echoes of the modern synthesis. Molecular Ecology 21: 4171–4189.

Kass, R.E., and Raftery, A.E. 1995. Bayes factors. J. Amer. Statist. Assoc. 90:773–795.

Knowles, L.L. 2000. Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. Evolution 54: 1337–1348.

Knowles, L.L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. Molecular Ecology 10: 691–701.

Kok, P.J.R., MacCulloch, R.D., Means, D.B., Roelants, K., Van Boxclaer, I., and Bossuyt, F., 2012. Low genetic diversity in tepui summit vertebrates. Current Biology 22: 589–590.

Korneliussen, T.S., Moltke, I., Albrechtsen, A., and Nielsen, R. 2013. Calculations of Tajima's D and other neutrality test statistics form low depth next-generation sequencing data. BMC Bioinformatics 14: 289.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., and Frost, S.D.W. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Molecular Biology and Evolution 23: 1891–1901.

Lampert, K.P., Rand, A.S., Mueller, U.G., and Ryan, M.J. 2003. Fine-scale genetic pattern and evidence for sex-biased dispersal in the túngara frog, *Physalaemus pustulosus*. Molecular Ecology 12: 3325–3334.

Lanfear, R., Calcott, B., Ho, S.Y.W., and Guindon S. 2012. Partition Finder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29: 1695–1701.

Leaché, A.D., and McGuire, J.A. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. Molecular Phylogenetics and Evolution 39: 638–644.

Leaché A.D., and Fujita, M.K., 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proceedings of the Royal Society B 277: 3071–3077.

Lemmon, E.M., Lemmon, A.R., and Cannatella, D.C. 2007. Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). Evolution 61: 2086–2103.

Luikart, G., England, P.R., Tallmon, D., Jordan, S., and Taberlet, P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics 4: 981–994.

MacCulloch, R.D., and Lathrop, A. 2002. Exceptional diversity of *Stefania* (Anura: Hylidae) on Mount Ayanganna, Guyana: three new species and new distribution records. Herpetologica 58: 327–346.

Maddison, D., and Maddison, W. 2001. MacClade 4: Analysis of phylogeny and character evolution. Version 4.08. Sinauer Associates, Sunderland, MA.

Maguire, B. 1970. On the flora of the Guayana Highland. Biotropica 2: 85–100.

Masta, S.E. 2000. Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): recent vicariance of sky island populations? Evolution 54: 1699–1711.

Mayr, E. and Phelps, W.H. 1955. The origin of the bird fauna of South Venezuelan Highlands. Bulletin of the American Museum of Natural History 136: 269–328.

Mayr, E., and Phelps, W.H. 1967. The origin of the bird fauna of the south Venezuelan highlands. Bulletin of the American Museum of Natural History 136: 273–327.

Mayr, E., and Diamond, J. 1976. Birds on islands in the sky: origin of the montane avifauna of Northern Melanesia. Proceedings of the National Academy of Sciences USA 73: 1765–1769.

Mayr, E. and R. J. O'Hara. 1986. The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. Evolution 40: 55–67.

McDiarmid, R.W., and Donnelly, M.A. 2005. The herpetofauna of the Guayana Highlands: amphibians and reptiles of the Lost World. In: Donnelly, M.A., Crother, B.I., Guyer, C., Wake, M.H., White, M.E. (Eds.), Ecology and Evolution in the Tropics: a Herpetological Perspective. University of Chicago Press, Chicago, Illinois, pp. 461–560.

Meikeljohn, C.D., Montooth, K.L., and Rand, D.M. 2007. Positive and negative selection on the mitochondrial genome. Trends in Genetics 23: 259–263.

Messer, P.W., and Petrov, D.A. 2013. Population genomics of rapid adaptation by soft selective sweeps. Trends in Ecology and Evolution 28: 659-669.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., and Johnson, E.A. 2007. Rapid and cost–effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17: 240–248.

Moen, D.S., and Wiens., J.J. 2009. Phylogenetic evidence for competitively driven divergence: Body-size evolution in Caribbean treefrogs (Hylidae: *Osteopilus*). Evolution 63: 195–214.

Moritz, C., Patton, J., Schneider, C., and Smith, T. 2000. Diversification of rainforest faunas: an integrated molecular approach. Annual Review of Ecology and Systematics 31: 533-563.

Naduvilezhath, L., Rose, L.E., and Metzler, D. 2011. Jaatha: a fast composite-likelihood approach to estimate demographic parameters. Molecular Ecology 20: 2709–2723.

Navas, C. A. 2006. Patterns of distribution of anurans in high Andean tropical elevations: insights from integrating biogeography and evolutionary physiology. Integrative and Comparative Biology 46: 82–91.

Noonan, B.P., and Gaucher, P. 2005. Phylogeography and demography of Guianan harlequin toads (*Atelopus*): diversification within a refuge. Molecular Ecology 14: 3017–3031.

Noonan, B. P. and Gaucher, P. 2006. Refugial isolation and secondary contact in the Dyeing Poison Frog *Dendrobates tinctorius*. Molecular Ecology 15: 4425–4435.

Nylander, J.A.A. 2004. MrModeltest v2. Evolutionary Biology Centre, Uppsala Univ.

Orme, C., Davies, R., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V., Webster, A., Ding, T., Rasmussen, P., and Ridgely, R. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436: 1016–1019.

Palo, J.U., Lesbarreres, D., Schmeller, D.S., Primmer, C.R., and Marilea, J. 2004. Microsatellite marker data suggest sex-biased dispersal in the common frog *Rana temporaria*. Molecular Ecology 13: 2865–2869.

Palumbi, S.R., Cipriano, F., and Hare, M.P. 2001. Predicting nuclear gene coalescence form mitochondrial data: the three-times rule. Evolution 55: 859–868.

Pauly, G., Hillis, D., and Cannatella, D. 2004. The history of a Nearctic colonization: molecular phylogenetics and biogeography of the nearctic toads (*Bufo*). Evolution 58: 2517–2535.

Pauly, G.B. 2008. Phylogenetic systematics, historical biogeography, and the evolution of vocalizations in the Nearctic toads (*Bufo*). PhD Dissertation, University of Texas, Austin.

Peterson, B., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. 2012. Double digest RADseq: an inexpensive method for De Novo SNP discovery and genotyping in model and non-model species. PLoS ONE doi: 10.1371/journal.pone.0037135

Posada, D. 2004. COLLAPSE 1.2. Available at http://darwin.uvigo.es/software/collapse.html.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Rambaut, A., and Drummond, A.J. 2013. Tracer v1.6 [Online]. Available: http://tree.bio.ed.ac.uk/software/tracer/

Rannala, B., and Yang, Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164: 1645–1656.

Raposo Do Amaral., F., Albers, P.K., Edwards, S.V., and Miyaki, C.Y. 2013. Multilocus tests of Pleistocene refugia and ancient divergence in a pair of Atlantic Forest antbirds (*Myrmeciza*). Molecular Ecology 22: 3996–4013.

Rato, C., Carranza, S., Perera, A., Carretero, M.A., and Harris, D.J. 2010. Conflicting patterns of nucleotide diversity between mtDNA and nDNA in the Moorish gecko, *Tarentola mauritanica*. Molecular Phylogenetics and Evolution 56: 962–971.

Robin, V.V., Sinha, A., and Ramakrishnan U. 2010. Ancient geographical gaps and paleo-climate shape the phylogeography of an endemic bird in the sky islands of southern India. PLoS ONE 5: e11321. doi:10.1371/journal.pone.0013321.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Ronquist F., Teslenko M., Van der Mark P., Ayres D., Darling A., Hohna S., Larget B., Liu L., Suchard M.A., and Huelsenbeck J.P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61: 539–542.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4: 137–138.

Rull, V. 2004a. An evaluation of the Lost World and Vertical Displacement hypotheses in the Chimantá Massif, Venezuelan Guayana. Global Ecology and Biogeography 13: 141–148.

Rull, V. 2004b. Is the "Lost World" really lost? Palaeoecological insights into the origin of the peculiar flora of the Guayana Highlands. Naturwissenschaften 91: 139-142.

Rull, V. 2004c. Biogeography of the Lost World, a palaeoecological perspective. Earth-Science Reviews 67: 125-137.

Rull, V. 2005. Biotic diversification in the Guayana Highlands: a proposal. Journal of Biogeography 32: 921–927.

Rull, V. 2008. Speciation timing and Neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. Molecular Ecology 17: 2722–2729.

Rull, V. and Nogué, S. 2007. Potential migration routes and barriers for vascular plants of the Neotropical Guyana Highlands during the Quaternary. Journal of Biogeography 34: 1327–1341.

Salducci, M.D., Marty, C., Chappaz, R., and Gilles, A. 2002. Molecular phylogeny of French Guiana Hylinae: implications for the systematic and biodiversity of the Neotropical frogs. Comptes Rendus Biologiques 325: 141–153.

Salerno, P.E. and Pauly, G.B. 2012. Clutch size variation in egg-brooding *Stefania*. South American Journal of Herpetology 7: 47–54.

Salerno, P.E., Ron, S.R., Señaris, J.C., Rojas-Runjaic, F.J.M., Noonan, B.P., and Cannatella, D.C. 2012. Ancient tepui summits harbor young rather than old lineages of endemic frogs. Evolution 66: 3000–3013.

Salerno, P.E., Señaris, J.C., Rojas-Runjaic, F.J.M. , and Cannatella, D.C. *In revision*. Recent evolutionary history of Lost World endemics: population genetics, species delimitation, and phylogeography in a highly complex sky-island landscape. Molecular Phylogenetics and Evolution.

Salerno, P.E., Rojas-Runjaic, F.J.M., Cannatella, D.C., and Señaris, J.C. *In preparation*. A new morphologically cryptic species of *Tepuihyla* from the sky islands of the Chimantá Massif in Venezuela. ZooKeys.

Sanchíz, B. 1998. Salientia. Vol. 4. Verlag F. Pfeil, München, Germany.

Seeb, J.E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., and Seeb, L.W. 2011. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11: 1–8.

Smith, C., and Farrell, B. 2005. Phylogeography of the longhorn cactus beetle *Moneilema appressum* LeConte (Coleoptera: Cerambycidae): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? Molecular Ecology 14: 3049–3065.

Smith, S.A., Stephens, P.R., and Wiens, J.J. 2005. Replicate patterns of species richness, historical biogeography, and phylogeny in Holarctic treefrogs. Evolution 59: 2433–2450.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

Stephens, M., and Donnelly, P. 2003. A comparison of bayesian methods for haplotype reconstruction. American Journal of Human Genetics 73: 1162–1169.

Stephens, M., Smith, N.J., and Donnelly, P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68: 978–989.

Steyermark, J.A. 1979. Flora of the Guayana Highland: Endemicity of the generic flora of the summits of the Venezuela tepuis. Taxon 28: 45–54.

Spear, S.F., Balkenhol, N., Fortin, M-J., McRae, B.H., and Scribner, K. 2010. Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Molecular Ecology 19: 3576–3591.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L., and Waits, L.P. 2007. Putting the 'landscape' in landscape genetics. Heredity 98: 128–142.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tamura, K.D., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731–2739.

Teacher, A.G.F., and Griffiths, D.J. 2011. HapStar: Automated haplotype network layout and visualisation. Molecular Ecology Resources 11: 151–153.

Theunert, C., Tang, K., Lachmann, M., Hu, S., and Stoneking, M. 2012. Inferring the history of population size change from genome-wide SNP data. Molecular Biology and Evolution 29: 3653–3667.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Resources 25: 4876–4882.

Trueb, L. 1970. Evolutionary relationships of casque-headed tree frogs with co-ossified skulls (family Hylidae). University of Kanssas Publications of the Museum of Natural History 18: 547–716.

Vences, M., Thomas, M., Van der Meijden, A., Chiari, Y., and Vieites, D. 2005. Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. Frontiers in Zoology 2: doi:10.1186/1742-9994-2-5.

Wiens, J.J., Kuczynski, C.A., Hua, X., and Moen, D.S. 2010. An expanded phylogeny of treefrogs (Hylidae) based on nuclear and mitochondrial sequence data. Molecular Phylogenetics and Evolution 55: 871–882.

Yang, Z., and Rannala, B. 2010. Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences USA 107: 9264–9269.